[i]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.

Thank you for the positive words.

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.

Response: Our goal with the sMLR model is indeed to gradually demonstrate that ANN is better than traditional linear/multilinear models. In the revised version, we added more in-depth discussion of each model results, and added more details about the multilinear model. We also tried to minimize any repetition of previous findings.

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.

Response: This is a good point. Thank you for pointing this out.
In the revised version, we followed the guideline introduced by Galí et al. 2015 to conduct quality control. Specifically, we removed DMS data with concentration less than 0.1 nM and greater than 100 nM, we also removed data with salinity less than 30 psu, so that we focus our

study on the open ocean. We removed DMSPt data that are less than 1 nM. Other than that, we did not do any binning and averaging to preserve the original data variance. Finally, there are 10404 pairs of DMS-Chl-*a* and 4061 pairs DMS-DMSPt, which is substantially more than what was reported by Galí et al. 2018 (with 3637 DMS-DMSPt and 8141 DMS-Chl pairs). This is because, the PMEL database has been greatly expanded, it now has over 80K DMS data points.

For Chl a data, we have added the following discussion (l.84 – 1.92).
"SeaWiFS Chl-*a* data (Level 3-binned, spatial resolution of 9.2 km) from December 1997 to March 2010 were matched to DMS data according to coordinates and sampling date. We compared PMEL in situ Chl *a* to SeaWiFS Chl *a*, which are well correlated on logarithmic scale (R^2 = 0.64) with a slope of 0.67 and an intercept of -0.06, $[\log(Chl_{SeaWiFS}) = 0.67\log(Chl_{in-situ}) - 0.01]$, which means that on logarithmic scale SeaWiFS Chl-*a* concentrations are on average ~30% lower than those of in situ Chl-*a* concentrations. This is possibly because SeaWiFS Chl *a* is calibrated based on HPLC determined Chl *a* (Morel et al., 2007), which on average is ~40% lower than that determined using Fluorometric method (Sathyendranath et al., 2009). Unfortunately, there is no flag in the database showing how Chl *a* was determined. For consistency, we use only Chl-*a* data retrieved from SeaWiFS in the following multilinear and network models."

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).

Response: Since our initial submission, the PMEL database dramatically expanded. Now there are a total of 86,785 valid DMS measurements (concentration greater than 0.1 nM and less than 100 nM according to your instructions), that is 71% larger than the number of data we initially used (51,161). For the expanded data set, ~93% of DMS are accompanied with in situ SST measurements, ~81% are accompanied with in situ salinity measurements. More importantly, each data point has their unique location and sampling time signatures. As shown in the following figure, sampling time (date) and location information is a strong DMS predictor, which together can decrease DMS root mean square error to 0.64 (on natural logarithm scale). Adding other climatological predictors can further improve the model performance.

The NAAMES dataset has 6,786 valid data points, which are ~7% of the total data points (93,571 = 86785+6786). All data are accompanied with in situ Chl *a*, SST, and SAL measurements. For parameters without in situ measurements, high resolution data are used to match DMS measurements, 0.0417° ×0.0417° for PAR, 0.5° ×0.5° for MLD, and 1° ×1° for NO3, which ensures most of DMS have a set of unique predictors. As shown in Table 1, merging NAAMES data with PMEL data does not significantly change the statistic.

Moreover, binning the data will reduce data variance, which has been demonstrated by Derevianko et al. (2009). The objective of this study is to train an ANN with as much data as possible, and let the ANN do any fitting. The statement "the final purpose is calculating a monthly climatology at coarser spatial resolution" is only partially true. The model can be applied to coarse resolution predictor fields, but also to very fine resolution predictor fields. For example, we have applied the network to fine resolution NAAMES fields for comparison with in situ DMS measurements (Bell et al., in preparation).

Lastly, binning data will necessarily result in loss of information. A great amount of information is associated with sampling location and date as shown in the following figure (Fig. 2a in MS). By binning the data into monthly 1°× 1° grid, the number of data points decreases significantly from 82,996 to only 9,018; sampling date feature (365) will be average to 12 months, and coordination combinations will be averaged from 87,332 × 87,332 to 180° × 360°, which represents a substantial loss of information. For ANN models, using less data points can lead to overfitting (See Fig. 2b).

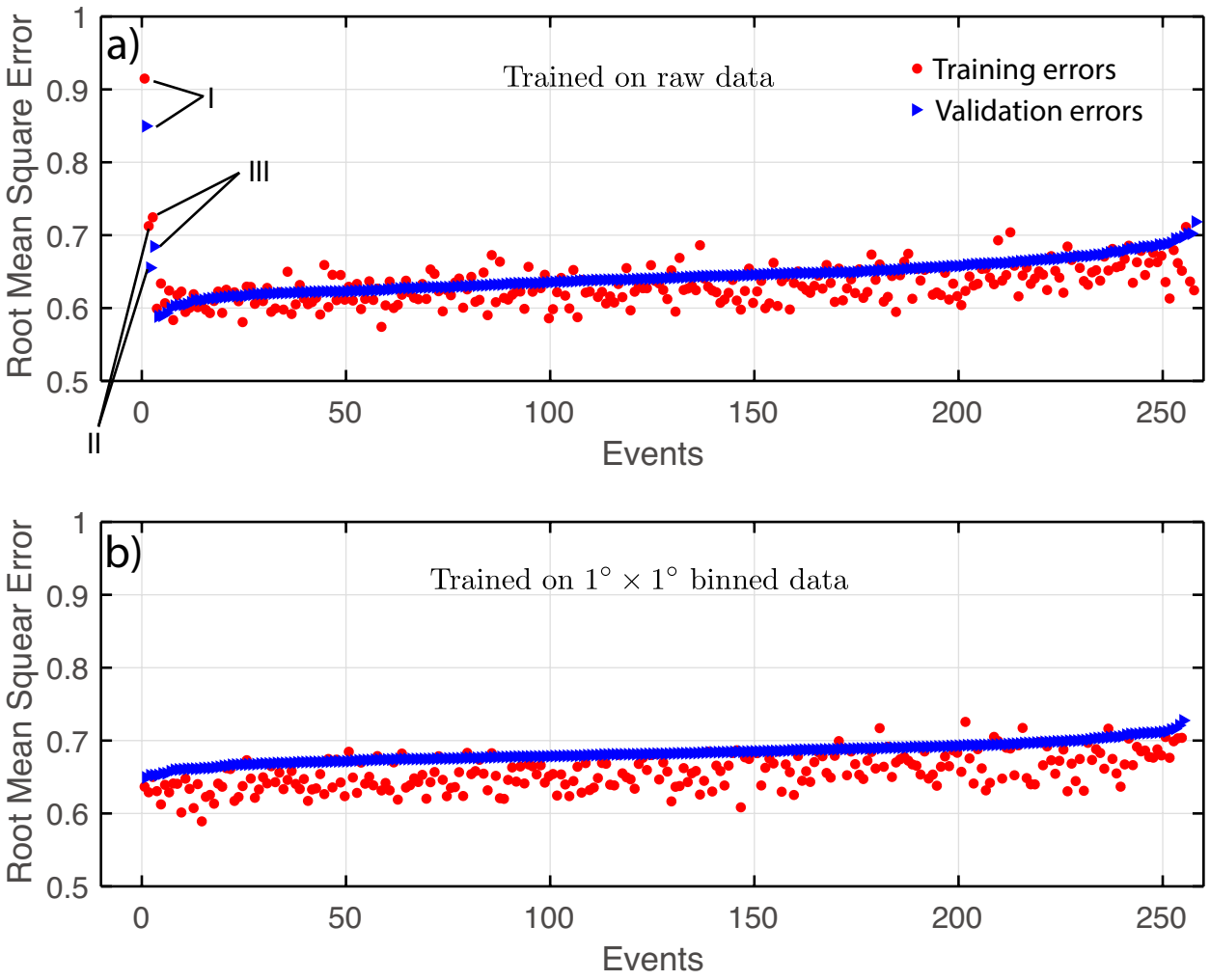

Fig. 2 Parameter sensitivity tests on raw and binned data. (a) Root mean square error on logarithmic scale for the model trained using raw data; (b) Root mean square error on logarithmic scale for the model trained using binned data. The time and location parameters are tested separately without combining with environmental parameters as shown in the upper panel, (I) with only location parameters; (II) with location and day of year parameters; and (III) with location, day of year, and time of day parameters. The model with three location parameters (I) has a root mean square error on natural logarithmic scale of ~0.83, which decreases to ~0.65 by adding sampling day of year parameters (II), however, increases to ~0.67 by adding sampling time parameters (III). We, therefore, do not include sampling time parameters in the following tests. We tested every possible combination of the eight parameters (PAR, MLD, SST, SAL, Chl a, DIP, DIN, and SiO), which in total are 255 tests.

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.

Response: We have added extensive discussion about the ANN approach on page 9 (l. 259 - l.274).

We also have added discussion emphasizing on comparison with previous models (P.12 l. 342 – l.364 and Figs. 3 and 4).

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.

Response: Thank you for the suggestions. We have updated our reference and revised the biogeochemical description of DMS/DMSP.

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.

This is an important point. It motivated us to do more tests to help open the “black box” as discussed in the revised MS. (l.259 – l. 264).

“From traditional linear or multilinear models, one can easily figure out which parameter is a strong predictor and how a predictor influence the state variable (e.g. the correlation between DMSP and DMS). An ANN model is much more complex, it adjusts weights of each node that connect inputs and outputs, therefore, the relationship between inputs and outputs is subtle. That’s why ANN models are generally referred to as a "Black Box". In this study, we design experiments that help open this "Black Box" and reveal parameters that drive surface ocean DMS distributions.”

And please also read the subsequent discussion (l.264 – l.317).

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).

Thank you for your constructive comments.
I hope you will find that the paper has been greatly improved based on comments of you and two other anonymous reviewers.

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.
We enjoyed reading your papers.

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).
We have taken most of your advice, and the manuscript has been thoroughly revised. More tests have been conducted, and our results do not significantly change, which means our method is robust. The abstract has been edited accordingly.

Introduction

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

We have added more tests to figure out how the environmental changes can influence surface ocean DMS distribution. With this revision, we believe that the sentence here is appropriate.

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.

The term “process rate measurements” has been added.
Meanwhile, we did more experiments to exploit the mechanisms.

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.

Corresponding reference has been added. Thank you.

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).

Corresponding references have been added. Thank you.

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.

Corresponding references have been added, and now the text is as follows,

“The precursor of DMS, DMSP, is mainly produced by marine algae (e.g. Kiene et al., 2000; Curson et al., 65 2011), and a small fraction of DMSP is transformed to DMS by marine algae and/or bacteria lyases (Simó, 2001; Stefels et al., 2007; Curson et al., 2011; Moran et al., 2012), and mostly as a result of food web interactions (Kiene et al., 2000; Simó, 2001).”

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.

The valid NAAMES data points are 6,939, which accounts for ~7% of the total data with expanded PMEL database. For the NAAMES data, we matched the DMS observation with super high-resolution satellite products, which ensures most of the data points have unique predictor combinations. We have added more results and discussion to argue why binning the data is not a good choice.

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.

Thank you for the useful tips, we followed your instruction to clean up the data. Specifically, we removed DMS concentration higher than 100 nM and lower than 0.1 nM, DMSPt concentration lower than 1 nM. We also removed data with salinity lower than 30 psu to focus on open ocean.

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).
Yes, Good point.
For Chl a data, we have added the following discussion (l.84 – 1.92).

"SeaWiFS Chl-*a* data (Level 3-binned, spatial resolution of 9.2 km) from December 1997 to March 2010 were matched to DMS data according to coordinates and sampling date. We compared PMEL in situ Chl *a* to SeaWiFS Chl *a*, which are well correlated on logarithmic scale (R^2 = 0.64) with a slope of 0.67 and an intercept of -0.06, $[\log(Chl_{SeaWiFS}) = 0.67\log(Chl_{in-situ}) - 0.01]$, which means that on logarithmic scale SeaWiFS Chl-*a* concentrations are on average ~30% lower than those of in situ Chl-*a* concentrations. This is possibly because SeaWiFS Chl *a* is calibrated based on HPLC determined Chl *a* (Morel et al., 2007), which on average is ~40% lower than that determined using Fluorometric method (Sathyendranath et al., 2009). Unfortunately, there is no flag in the database showing how Chl *a* was determined. For consistency, we use only Chl-*a* data retrieved from SeaWiFS in the following multilinear and network models."

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.
Good point. We updated our climatology so that the current climatology used is from 1997-2010. The new climatology is Level 3-binned (L3BIN with spatial resolution of 9.2 km) from SeaWiFS.
The change of Chl a climatology does not significantly change our results.

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.
Good point. We updated our MLD climatology and used MIMMOC one in the revised model.
The change of MLD climatology does not significantly change our results.

L81: Please specify the nutrient datasets used.
Good point. We added descriptions about the nutrient data sources (l.98 – l.101).

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?
Good point, but it is partially true. Except for SST that has the most in situ observational data (81069 for PMEL data), in situ salinity was reported for ~74% of the DMS data. More importantly, every data point has their unique time-space signatures (5 parameters in the

model). We used high resolution Chl *a* (0.418), PAR (0.418), and MLD (0.5×0.5) climatologies for the PMEL data. For the NAMMS data, the Chl *a*, SST, and PAR have even higher resolution (0.0417×0.0417). The high-resolution data and unique time-space ensure that each data has a unique signature. See also our response to your general question No. 3.

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

These parameters are called hyper-parameters in machine learning language. We manually tuned the parameters to prevent over-/under- fitting the data (l.159 – l.163).

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).
Good point, we retested the diurnal parameters (two time parameters). Adding them slightly worsen the performance of the model (Fig. 2a in the MS). We have added the following discussion (l.265 – l.274),

"Given the strong correlation between solar radiation and DMS concentration reported by Vallina and Simó (2007), one would expect that adding sampling time would improve the model performance. However, it increases RMSE slightly (Fig. 2a). Galí et al. (2013c) studied diel cycle at the Mediterranean Sea and Sargasso Sea. Among their four experiments (three in the Mediterranean Sea and one in the Sargasso Sea) regular diel variation was observed at only one experiment in the Mediterranean Sea at summer season, with highest DMS values observed at midnight and lowest values at midday. In all the other experiments, diel variations for both DMS and DMSPt pools were small. Gross community DMS production during the daytime was two to three times higher than that in the nighttime, but the high DMS production was compensated by greater photochemical and microbial consumption (Galí et al., 2013c). The balance between DMS production and consumption appears to dampens DMS diel variation. This may explain why adding time parameters does not improve the ANN model's predictive ability."

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).

We followed your instructions to clean up the data. Our new R^2 value for DMS and in situ Chl *a* is 0.21 (n = 10,404 compared to 8141 in Gali et al. (2018)), for DMS and DMSPt is 0.41 (n = 4060, compared to 3637 in Gali et al. (2018)). Both pairs have larger data than those by Gali et al. (2018) for two reasons, 1) the expansion of the online database; 2) no averaging being done.

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.

We have removed the corresponding sentence. Meanwhile, we have added more discussion as follows (l.226 – l.229).

"McParland and Levine (2019) developed a mechanistic model that related intracellular DMSP concentration to environmental stress, and coupled the model with MIT ecosystem model (DARWIN) to estimate global ocean DMSP distribution. Galí et al. (2015) first applied a remote sensing algorithm to obtain a DMSP climatology, from which they predict DMS climatology through an empirical relationship with PAR (Galí et al., 2018)."

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.
Thank you for pointing this out.
We made the following corrections (l.314 – l.317).

"On the other hand, there are numerous studies that observed no correlation between DMS and Chl a (e.g. Dacey et al., 1998; Kettle et al., 1999; Toole and Siegel, 2004). The inconsistent relationships indicate the complexity of the biogeochemical reduced sulfur cycle. As suggested by Simó (2001), not only can phytoplankton biomass, taxonomy, and activity influence DMS production, but so does food-web structure and dynamics. The inconsistent relationship may also explain the low ranking of Chl a in the models."

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.
Thank you for pointing this out.
We have clarified the reasoning in the text and also as follows (l.289 – l.296).

"MLD is another important predictor. High DMS concentrations in the open ocean have been detected when the water column is most stratified (Simó and Pedrós-Alió, 1999). The authors proposed that a stratified (high light) environment nourishes strong DMSP producers, or that phytoplankton cellular DMSP quota increases in such an environment. High conversion rates from DMSP to DMS in stratified waters is another reason for high DMS concentrations when MLD is shallow. Meanwhile, the biological DMS consumption rate decreases in oligotrophic oceans (Galí and Simó, 2015). A dilution model was also proposed to explain the anti-correlation between DMS concentration and MLD (Aranami and Tsunogai, 2004). The authors proposed that mixed layer deepening entrains water with little or no DMS into surface waters and dilutes surface DMS concentrations, but recent studies have shown that DMS loss rate via vertical mixing is orders of magnitude lower than production/consumption rates (e.g. Galí et al., 2013c; Royer et al., 2016)."

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.
We have updated the reference. Thank you.

Section 3.2: see general comment 3 on data binning.
Please refer to our response to your comments #2.

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?
Yes, we used exactly the same model, and have made this clear in the text as follows (l.319 – l.329),

"Simó and Dachs (2002) obtained high R^2 values between DMS concentration and the ratio of Chl a to MLD (Chl/MLD) when Chl/MLD is greater than or equal to 0.02, and between DMS concentration and ln(MLD) when Chl/MLD is less than 0.02. We tried exactly the same model on raw PMEL data with in situ Chl-a measurements and climatological MLD, and found that both correlations between DMS and Chl/MLD (n = 4,921, R^2 =~ 0 .1) and between DMS and ln(MLD) (n = 5,978, R^2 =~ 0) are statistically insignificant. To reduce interannual variability, we binned in situ Chl a and DMS into monthly 1° × 1° grid, and retested the above model. We found that the correlations are still statistically insignificant. (R^2=~ 0)

Vallina and Simó (2007) reported an R2 of 0.95 (n=14) between DMS concentration and SRD. We applied the same linear regressions on both raw data and monthly 1° × 1° data, and found

no significant correlations between DMS and SRD as calculated according to Vallina and Simó (2007):,

$$SRD = SI \cdot \frac{1}{Kd490 \cdot MLD}(1 - \exp(-Kd490 \cdot MLD),$$

where SI is solar insolation (W m-2); Monthly SI data are from a CESM simulation (Wang et al., 2019)."

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?

The Choice of parameter combinations are based on BIC criterion.
Yes, R^2 value is for natural log space. We have made this clear in the revised paper.

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.

Yes, we have extensively discussed the ANN model on pages 9.

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.

Good resources. We have plotted the data together with our new prediction (Fig.3 and Fig. 4).

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.

Good point, we updated the discussion as follows (l.371 – l.376),

"The summertime high DMS concentration at high latitudes is believed to be linked to the release of ice algae that are prolific DMSP producers (Stefels et al., 2012; Webb et al., 2019). As an important cryoprotectant and osmolyte, DMSP helps ice algae to cope with the low temperature and high salinity conditions (Thomas and Dieckmann, 2002). High DMS concentrations at high latitudes have also been observed to accompany blooms of coccolithophores and Phaeocystis, which are strong DMSP producers (Neukermans et al., 2018; Wang et al., 2015). The shoaling of mixed layer depth in summer provides favorable conditions, i.e. stable and warm, with adequate irradiation for coccolithophores and Phaeocystis growth (Galí et al., 2019)"

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).

In the revised MS, we rerun the model, and now there is not region with DMS concentration lower than 0.1 nM.

L250-254: Please check Galí and Simó 2015 (GBC) for a mechanistic explanation of the summer paradox.
We have changed the wording as follows (l.389 – l.393),

"Fig. 6 compares monthly mean Chl-a concentrations to DMS concentrations in N. and S. hemisphere gyres. The concentrations are normalized to the range of 0 to 1. It is clear that Chl a and DMS are anti-correlated, DMS concentration peaks at summer season when Ch-a concentration is generally low. This phenomenon is previously termed as "summer DMS paradox" (Simó and Pedrós-Alió, 1999). This pattern is more apparent in the S. hemisphere gyres, because the terrestrial influence is smaller in the S. hemisphere than in the N. hemisphere."
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.

Good point.
We plotted monthly mean Chl-*a* and DMS concentrations for N. and S. hemisphere gyres as shown in the following figures. It is clear that DMS and Chl *a* are anti-correlated in the gyres. DMS peaks in the summer when Chl *a* is at annual minimum.

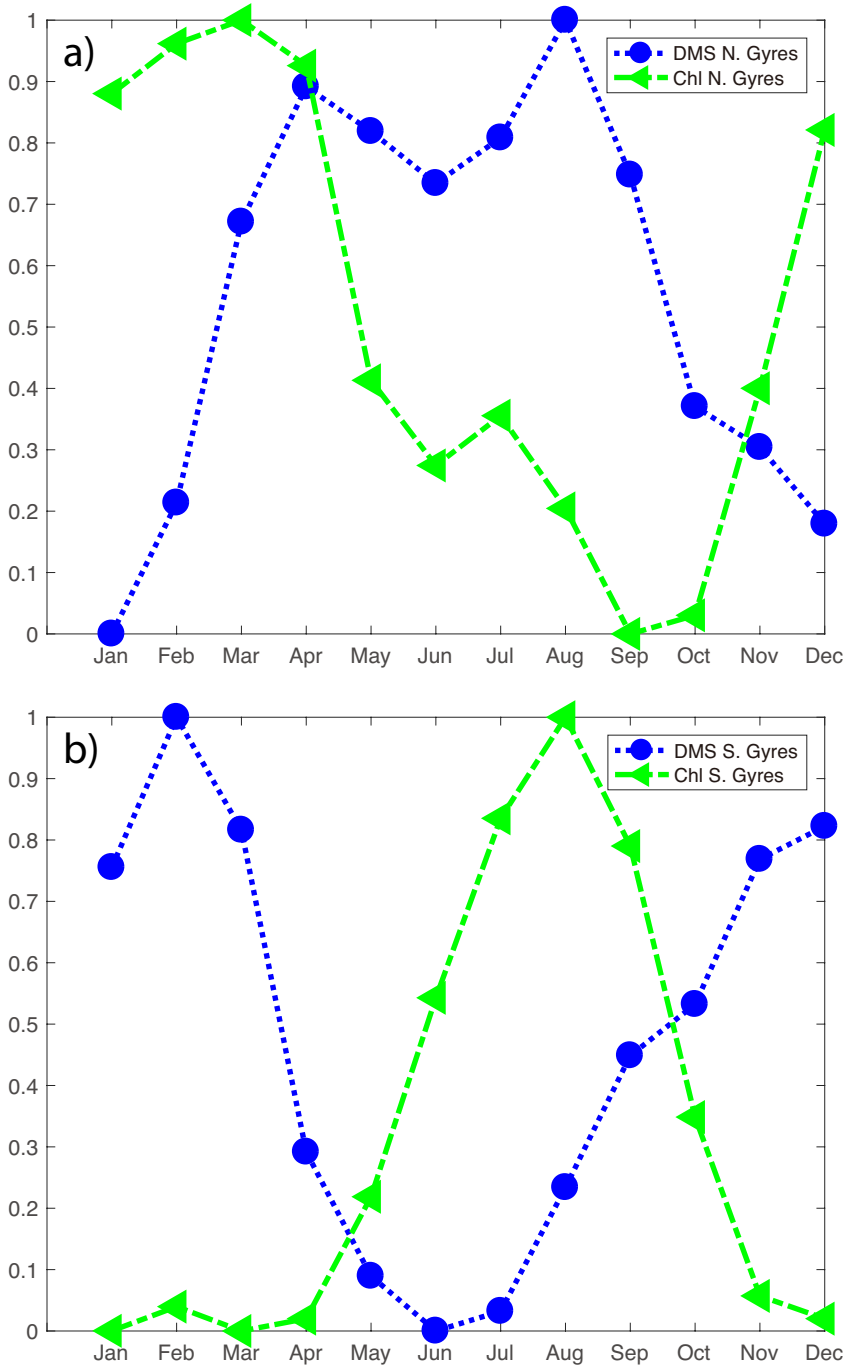

Figure 8. Distributions of monthly mean DMS and Chl-*a* concentrations for N. and S. hemisphere gyres. The gyres are defined as regions between 30◦ and equator where annually mean DIP concentration is below 0.2 µM. Monthly mean concentrations are normalized to the range of 0 to 1.

L273-293: Here I strongly suggest citing Tesdal et al. 2016.
Reference has been cited, and more discussion has been added.

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

Thank you for this suggestion, we have updated the colormaps for both figures.

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.

Yes, another good suggestion, we updated the figure accordingly.

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

one **epoch** = one forward pass and one backward pass of *all* the training examples

In terms of artificial neural networks, an epoch refers to one cycle through the full training dataset. Usually, training a neural network takes more than a few epochs. In other words, if we feed a neural network the training data for more than one epoch in different patterns, we hope for a better generalization when given a new "unseen" input (test data).

We have added more explaination in the revised MS (l.167 – l.169),
An epoch consists of one forward pass and one backward pass of all the training examples.

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

Thank you. We have removed the term.

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.

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 con- struct 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 accepted 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.

Thank you for your positive comments.

The text is generally well written and the display 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 para- dox”, 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.

Thank you for your correction.

We have rephrased the statement and added corresponding reference as follows,

“The weak relationship may be caused by the so-called “summer DMS paradox”, which describes a phenomenon that annual maximum of surface DMS concentration is commonly

detected in summer when Chl *a* is at its annual minimum in mid and subtropical low latitude waters (Simó and Pedrós-Alió, 1999)."

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.

We have corrected the corresponding statement as follows:

"Simó and Dachs (2002) achieved a strong linear relationship between heavily binned/averaged DMS and mixed layer depth (MLD) when Chl-a/MLD ≥ 0.02, and a logarithmic relationship between DMS and Chl-a/MLD when Chl-a/MLD < 0.02."

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.

Thank you for pointing this out. We have adjusted the description as follows,

"Many provinces lacked adequate data to create a reliable climatology (Fig. A1). In those situations, they first generated an annual cycle with monthly means for each province. Temporal interpolations were used to fill the monthly gaps if there were enough data to create a robust annual mean. Otherwise, interpolation from neighboring provinces was used to fill the remaining gaps."

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.

We have rephrased the statement as follows,

"The precursor of DMS, DMSP, is mainly produced by marine algae (e.g. Kiene et al., 2000; Curson et al., 2011), and a small fraction of DMSP is transformed to DMS by marine algae and/or bacteria lyases (Simó, 2001; Stefels et al., 2007; Curson et al., 2011; Moran et al., 2012), and mostly as a result of food web interactions (Kiene et al., 2000; Simó, 2001)."

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.

We have re-run the model using log-transformed K.

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.

Thank you for pointing this out. In the revised MS, we used only satellite Chl-*a*, and added more discussion as follows (l.84 – 1.92),

"SeaWiFS Chl-*a* data (Level 3-binned, spatial resolution of 9.2 km) from December 1997 to March 2010 were matched to DMS data according to coordinates and sampling date. We compared PMEL in situ Chl *a* to SeaWiFS Chl *a*, which are well correlated on logarithmic scale (R^2 = 0.64) with a slope of 0.67 and an intercept of -0.06, $[\log(Chl_{SeaWiFS}) = 0.67\log(Chl_{in-situ}) - 0.01]$, which means that on logarithmic scale SeaWiFS Chl-*a* concentrations are on average ~30% lower than those of in situ Chl-*a* concentrations. This is possibly because SeaWiFS Chl *a* is calibrated based on HPLC determined Chl *a* (Morel et al., 2007), which on average is ~40% lower than that determined using Fluorometric method (Sathyendranath et al., 2009). Unfortunately, there is no flag in the database showing how Chl *a* was determined. For consistency, we use only Chl-*a* data retrieved from SeaWiFS in the following multilinear and network models."

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?

In the paper we used two flux parameterizations, GM12 according to Goddijn-Murphy et al .2012 and N00 according to Nightingale et al 2000. GM12 DOES describe a linear relationship between Kw and u10, and it is more updated that Marandino et al 2008 (if this is the reference that you referred to).

We also used N00, because we can compare to previous results that used the same parameterization.

We did a correction on mean u10 as described below (l.209 – l.214),

"Because the N00 parameterization was calibrated using in situ wind speeds and has a nonlinear quadratic dependence on wind speed, the use of monthly mean wind speeds will introduce errors. To reconcile differences between in situ wind speed and monthly mean wind speed, a correction is applied according to Simó and Dachs (2002) by assuming that instantaneous wind speeds follow a Rayleigh distribution. Eq. 8 thus becomes $k_{w,660}$ $=[0.222\eta^2\Gamma(1+2/\xi)+0.333\eta\Gamma(s)](ScDMS/600)^{-0.5}$, where $\eta^2 = 4U_{10}^2/\pi$; $s=(1+1/\xi)$, and $\xi = 2$ for Rayleigh distribution (Livingstone and Imboden, 1993)."

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

We have added the following discussion (l.226 – l.229),

"McParland and Levine (2019) developed a mechanistic model that related intracellular DMSP concentration to environmental stress, and coupled the model with MIT ecosystem model (DARWIN) to estimate global ocean DMSP distribution. Galí et al. (2015) first applied a remote sensing algorithm to obtain a DMSP climatology, from which they predict DMS climatology through an empirical relationship with PAR (Galí et al., 2018)."

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.

Thank you for pointing this out.
We realized that it is that Dacey et al. (1998) reported the original Sargasso Sea data, and Toole and Siegel analyzed the correlation between DMS and Chl a.
We have added the original reference and changed the wording.

"On the other hand, there are numerous studies that failed to correlate DMS and Chl a (e.g. Dacey et al., 1998; Kettle et al., 1999; Toole and Siegel, 2004)."

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 = a*Ln(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.

Thank you for pointing this out. We have added more discussion as follows,

"It is proposed that stratified (high light) environment nourishes strong DMSP producers, or phytoplankton cellular DMSP quota increases in such an environment. High conversion rate from DMSP to DMS in stratified waters is another reason for high DMS concentrations when MLD is shallow. Meanwhile, the biological DMS consumption rate decreases in oligotrophic oceans (Galí and Simó, 2015). Dilution model, which describes a phenomenon that when the mixed layer deepens water with no or little DMS is entrained into the surface waters and dilutes surface DMS concentrations, was also proposed to explain the anti-correlation between DMS concentration and MLD (Aranami and Tsunogai, 2004). However, recent studies show that DMS loss rate via vertical mixing is orders of magnitude lower than production/consumption rates (e.g. Galí et al., 2013c; Royer et al., 2016)."

L189-199: There are a number of papers that should be cited here – besides Toole et al. and Sunda et al, several 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.

Good points.
We have rebuilt the SRD vs DMS model, recomputed the results, and revised our discussion as follows,

"Vallina and Simó (2007) reported an R^2 of 0.95 (n=14) between DMS concentration and SRD. We applied the same linear regressions on both raw data and monthly 1°× 1° data, and found no significant correlations between DMS and SRD that is calculated according to Vallina and Simó (2007) using the following equation,

$$SRD = SI \cdot \frac{1}{Kd490 \cdot MLD}(1 - \exp(-Kd490 \cdot MLD))$$

where SI is shortwave irradiance (W m^{-2}), which is converted from PAR according to Galí and Simó (2015)."
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.
Yes, we used exactly the same model, and have made it clear in the text as follows (l.319 – l.329),

"Simó and Dachs (2002) obtained high R^2 values between DMS concentration and the ratio of Chl a to MLD (Chl/MLD) when Chl/MLD is greater than or equal to 0.02, and between DMS concentration and ln(MLD) when Chl/MLD is less than 0.02. We tried exactly the same model on raw PMEL data with in situ Chl-a measurements and climatological MLD, and found that both correlations between DMS and Chl/MLD (n = 4,921, R^2 =~ 0 .1) and between DMS and ln(MLD) (n = 5,978, R^2 =~ 0) are statistically insignificant. To reduce interannual variability, we binned in situ Chl a and DMS into monthly 1° × 1° grid, and retested the above model. We found that the correlations are still statistically insignificant. (R^2=~ 0)"
."

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 results were recomputed and the text was rewritten. Please see our responses to your above comments.

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.

Since this issue was brought up by all three reviewers, we have added adequate discussion/arguments to this point as shown in the following text.

The PMEL database expanded dramatically. Now there are a total of 86,785 valid DMS measurements (concentration greater than 0.1 nM and less than 100 nM according to your instructions), that is 71% larger than the number of data we initially used (51,161). For the expanded data set, ~93% of DMS are accompanied with in-situ SST measurements, ~81% are accompanied with in-situ salinity measurements. More importantly, each data point has their unique location and sampling time signatures. As shown in the following figure, sampling time (date) and location information is a strong DMS predictor, which together can decrease DMS root mean square error to 0.64 (on natural logarithm scale). Adding other climatological predictors can further improve the model performance.

The NAAMES dataset has 6,786 valid data points, which are ~7% of the total data points (93,571 = 86785+6786). All data are accompanied with in-situ Chl a, SST, and SAL measurements. For parameters without in-situ measurements, high resolution data are used to match DMS measurements, 0.0417° ×0.0417° for PAR, 0.5° ×0.5° for MLD, and 1° ×1° for NO3, which ensures most of DMS have a set of unique predictors. As shown in Table 1, merging NAAMES data with PMEL data does not significantly change the statistic.

Moreover, binning the data will reduce data variance, which has been demonstrated by Derevianko et al. (2009). The objective of this study is to train an ANN with as many data as possible, so that the model is generalized. It not only can apply to coarse resolution predictor fields, but also can apply to very fine resolution field, for example, we have applied the network to fine resolution NAAMES fields for comparison with in-situ DMS measurements (Bell et al., in prep.).

Lastly, binning data will also result in loss of information. A great amount of information is associated with sampling time and date as shown in the following figure (Fig. 2a in MS). By binning the data into monthly 1° × 1° grid, the valid DMS data points will decrease significantly from 82,996 to 9,018; sampling date feature (365) will be average to 12 months, and coordination combinations will be averaged from 87,332×87,332 to 180°×360°, which

represents great information reduction. For ANN models, less data points usually lead to overfitting Fig. 2b.

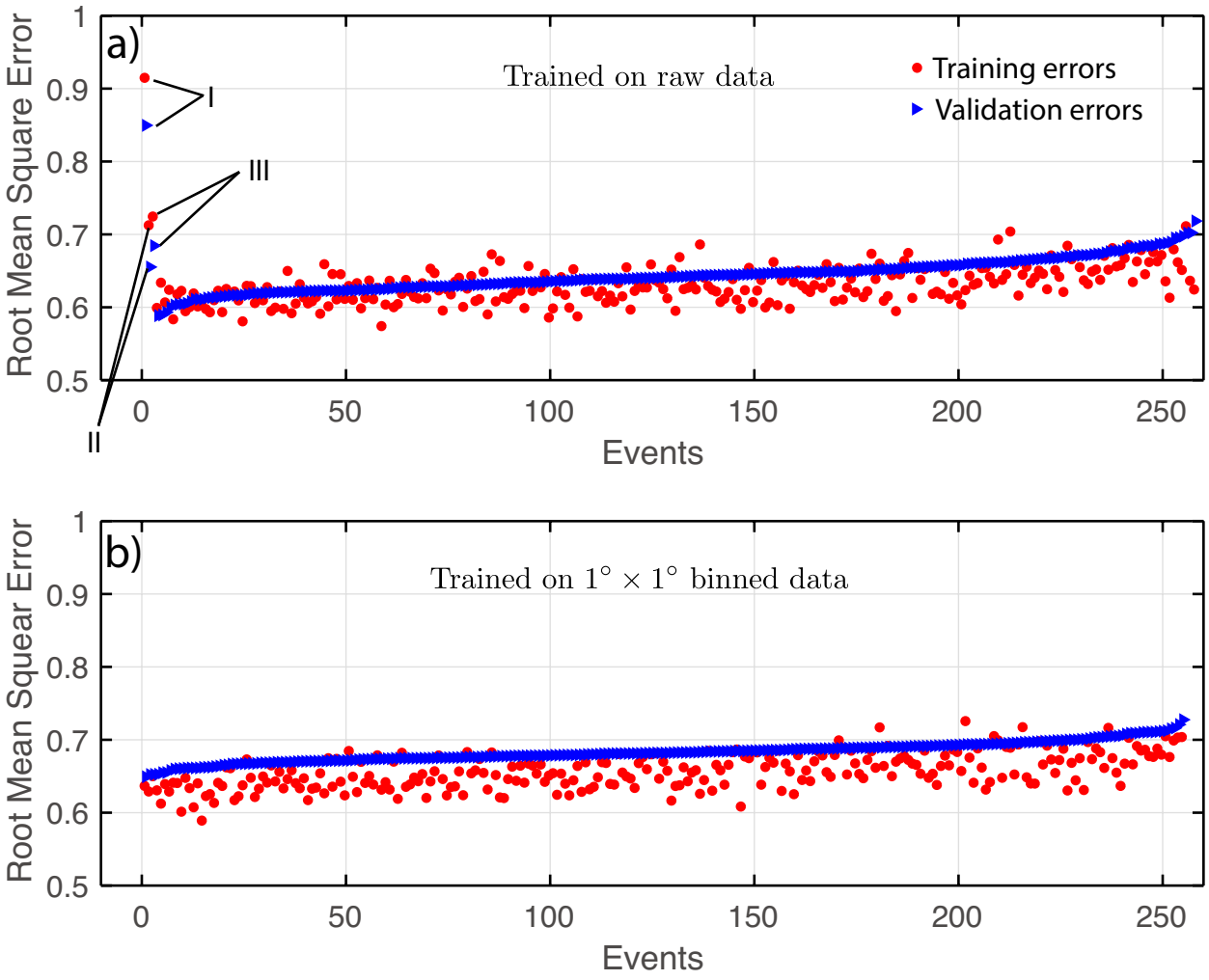

Fig. 2 Parameter sensitivity tests on raw and binned data. (a) Root mean square error on logarithmic scale for the model trained using raw data; (b) Root mean square error on logarithmic scale for the model trained using binned data. The time and location parameters are tested separately without combining with environmental parameters as shown in the upper panel, (I) with only location parameters; (II) with location and day of year parameters; and (III) with location, day of year, and time of day parameters. The model with three location parameters (I) has a root mean square error on natural logarithmic scale of ~0.83, which decreases to ~0.65 by adding sampling day of year parameters (II), however, increases to ~0.67 by adding sampling time parameters (III). We, therefore, do not include sampling time parameters in the following tests. We tested every possible combination of the eight parameters (PAR, MLD, SST, SAL, Chl a, DIP, DIN, and SiO), which in total are 255 tests.

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.
Good points. We have updated the reference, and added more discussion as follows,

Good point. We have updated the reference, and added more discussion as follows (l.371 – l.376),

“The summertime high DMS concentration at high latitudes is believed to be linked to the release of ice algae that are prolific DMSP producers (Stefels et al., 2012; Webb et al., 2019). As an important cryoprotectant and osmolyte, DMSP helps ice algae to cope with the low temperature and high salinity conditions (Thomas and Dieckmann, 2002). High DMS concentrations at high latitudes have also been observed to accompany blooms of coccolithophores and Phaeocystis, which are strong DMSP producers (Neukermans et al., 2018; Wang et al., 2015). The shoaling of mixed layer depth in summer provides favorable conditions, i.e. stable and warm, with adequate irradiation for coccolithophores and Phaeocystis growth (Galí et al., 2019)”

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.

In the revised MS, we have shortened the spatial distribution discussion, and added more discussion on the comparison with previous results.

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.

We agree that the provinces are defined from environmental descriptors, however, we also noticed that the provinces are static with no seasonality. In the revised version, we weakened our expression as follows,

“By contrast, objective interpolation methods are spatial/temporal averages of sparse data with weak underlying basis in environmental variability.”

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.

Good point. We split the global map into Southern and Northern hemispheres, and plotted seasonal cycles for each hemisphere (Fig. 3). Also, we made zonally mean average for each season (Fig. 4).

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.

This is an interesting point.

Based on Fig. A1, the observational data do not show high DMS hot spots/or good coverage in September in the subarctic NE Pacific, instead, there are some high DMS measurements in August in this region. Both Lana et al, 2011 and our ANN model capture the high DMS concentrations in August in this region.

We believe that in September the high DMS concentrations in this region are interpolated from August. On the other hand, our ANN model predicts moderate DMS concentration at this region in September.

In any event, we agree that more discussion is needed here. The following discussion was added to address the methodology difference between objective interpolation method and ANN method (l.353 – l.364).

"L11 stands out in the S. hemisphere monthly mean plot (Fig. 3b), with the highest mean concentrations in January and December, when DMS concentrations are ~2 times higher than other model predictions. Galí et al. (2018) identified five short-comings associated with the direct interpolation method employed by Lana et al. (2011). All shortcomings concern the nature of in situ DMS data, including right-skewed distribution, lack of spatial and temporal coverage, lack of duplicate measure- ments, and sampling bias towards DMS-productive conditions. Because of the sparsity and skewed distribution, the interpo- lation/extrapolation method broadcasts small scale features to large scales (Tesdal et al., 2016). This is especially true for the month of January and December when the elevated L11 monthly means were mainly driven by a small amount of extremely high DMS measurements (>40 nM) near the Antarctic continent. On the other hand, empirical models including the ANN model used in this study rely on environmental parameter climatologies to obtain the DMS climatology. Extreme conditions are smoothed out in climatological data, e.g. in the DMS database the maximum in situ Chl-a concentration is > 800 mg/m3, whereas it is ~50 mg/m3 in the SeaWiFS climatology. When climatological data are used to generate DMS distribution, a smaller variance than in situ data is expected."

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.

In the revised MS, we have added more tests and more discussion accordingly. More specifically, we designed experiments that help to open the network “black box”, and based on the results, we thoroughly discussed the parameters that exert an impact of the prediction ability of the model. The text is from line 253-line 303.

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 between 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.

Thank you for your positive comments.

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.

This point was also raised by the other two reviewers. We therefore dealt with it very carefully, and added the following arguments.

The PMEL database expanded dramatically. Now there are a total of 86,785 valid DMS measurements (concentration greater than 0.1 nM and less than 100 nM according to your instructions), that is 71% larger than the number of data we initially used (51,161). For the

expanded data set, ~93% of DMS are accompanied with in-situ SST measurements, ~81% are accompanied with in-situ salinity measurements. More importantly, each data point has their unique location and sampling time signatures. As shown in the following figure, sampling time (date) and location information is a strong DMS predictor, which together can decrease DMS root mean square error to 0.64 (on natural logarithm scale). Adding other climatological predictors can further improve the model performance.

The NAAMES dataset has 6,786 valid data points, which are ~7% of the total data points (93,571 = 86785+6786). All data are accompanied with in-situ Chl a, SST, and SAL measurements. For parameters without in-situ measurements, high resolution data are used to match DMS measurements, 0.0417° ×0.0417° for PAR, 0.5° ×0.5° for MLD, and 1° ×1° for NO3, which ensures most of DMS have a set of unique predictors. As shown in Table 1, merging NAAMES data with PMEL data does not significantly change the statistic.

Moreover, binning the data will reduce data variance, which has been demonstrated by Derevianko et al. (2009). The objective of this study is to train an ANN with as many data as possible, so that the model is generalized. It not only can apply to coarse resolution predictor fields, but also can apply to very fine resolution field, for example, we have applied the network to fine resolution NAAMES fields for comparison with in-situ DMS measurements (Bell et al., in prep.).

Lastly, binning data will also result in loss of information. A great amount of information is associated with sampling time and date as shown in the following figure (Fig. 2a in MS). By binning the data into monthly 1° × 1° grid, the valid DMS data points will decrease significantly from 82,996 to 9,018; sampling date feature (365) will be average to 12 months, and coordination combinations will be averaged from 87,332×87,332 to 180°×360°, which represents great information reduction. For ANN models, less data points usually lead to overfitting Fig. 2b.

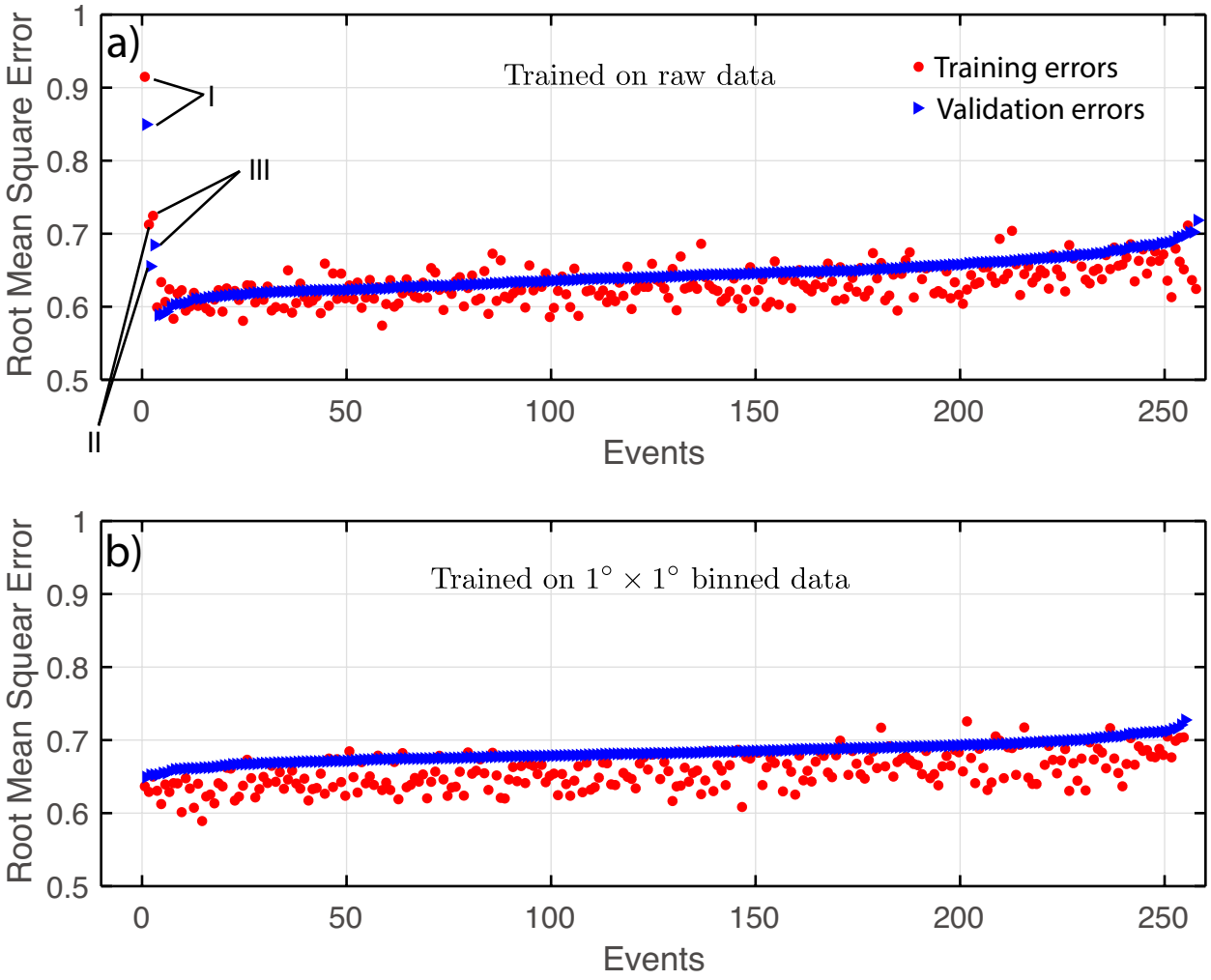

Fig. 2 Parameter sensitivity tests on raw and binned data. (a) Root mean square error on logarithmic scale for the model trained using raw data; (b) Root mean square error on logarithmic scale for the model trained using binned data. The time and location parameters are tested separately without combining with environmental parameters as shown in the upper panel, (I) with only location parameters; (II) with location and day of year parameters; and (III) with location, day of year, and time of day parameters. The model with three location parameters (I) has a root mean square error on natural logarithmic scale of ~0.83, which decreases to ~0.65 by adding sampling day of year parameters (II), however, increases to ~0.67 by adding sampling time parameters (III). We, therefore, do not include sampling time parameters in the following tests. We tested every possible combination of the eight parameters (PAR, MLD, SST, SAL, Chl a, DIP, DIN, and SiO), which in total are 255 tests.

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 crossvalidation 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 multiround 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.

Good point.
There are two general guidelines when one separates the data to training and validating sets, representation and generalization. That is to say that your training data has to be representative, and your model has to have the ability to generalize. The online DMS data are organized by contributor ID, while when you do cross validations, the data are drawn section by section as the following figure shows. One section of data may be from a specific contributor who collected data from a specific region, therefore, the data may not be representative, which results in an over-trained or a less-trained model (we have uploaded the cross-validation model to github ((https://github.com/weileiw/ANN-DMS-code), so interested readers can play with it.). To make the selection more representative, a common practice is to shuffle the data, and then randomly draw a fraction from the shuffled data. For DMS data (or maybe other oceanography data too), data collected from the same cruise are highly intercorrelated, so that shuffling and randomly splitting will "leak" information to the model and cause an overfitting (we have tested shuffling and random drawing method, it indeed leads to overfitting. (Code is also available at Github directory.)) .

Another purpose of doing cross-validation is to allow your model to see as many data as possible. This is useful when you do not have enough data to train your network. To achieve a similar effect, we first manually adjust the hyper-parameters (dropout ratio, hidden layers, number of nodes etc., they are key parameters to determine the model performance) using manually-divided training, internal testing, and external validation data. After we get a satisfactory combination of those hyper-parameters, we fix them and fine tune the network using all available data (because the data are intercorrelated, shuffle and randomly split training and testing does the work.).

Lastly, in the parameter selection experiments, we examined a total of 255 models (every combination of eight environmental parameters). We then ranked the model according to root mean squared error (RMSE) on validation data as shown in Fig. 2a. Compared to RMSEs on the training data, there are no apparent overfittings for the top 10 models. The models with larger RMSEs generally overfit the training data. Meanwhile, overfitting occurs with almost every model when binned data are used (Fig.2b).

Accordingly, we have added more explanations in the revised MS as follows (l.154 – l.162):

"The data was split into sets manually rather than automatically. The online DMS data are organized by contributor ID, and automatic splitting draws a continuous portion from the data. The data portion may come from a specific contributor who collected data from a specific region and it may therefore not be representative. This would result in an over-trained or a under-trained model. To make the selected data more representative, a common practice is to shuffle the data, and then randomly draw a fraction from the shuffled data. For DMS, data collected from the same cruise are highly intercorrelated, so that shuffling and randomly splitting "leaks" information to the model and causes overfitting. We manually adjust the hyper-parameters (dropout ratio, hidden layers, number of nodes etc.) using the data that has been manually-divided into training, internal testing, and external validation subsets. After obtaining a satisfactory combination of those hyper-parameters (as discussed below), we fix them and fine tune the network using all available data."

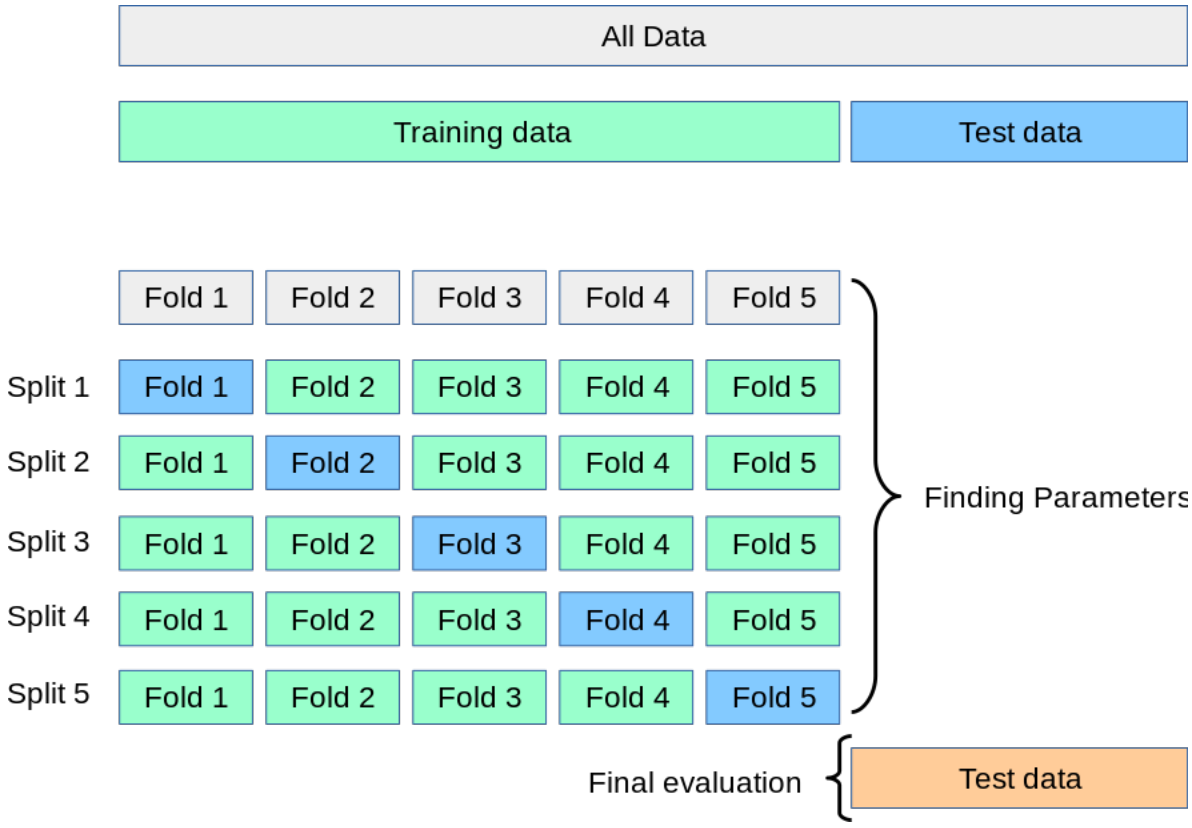

Figure cited from https://scikit-learn.org/stable/modules/cross_validation.html.

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?

This is a good point and following is the explanation.
In the ANN, there are at least two places using random states, 1) it uses random state to decide the Dropout nodes, 2) it uses random state to separate internal testing data from training data. The random states do not control initial parameter values, but different random states produce slightly different results. To make our model results reproduceable, we fixed the random state at 64 in the revised model. The uncertainly analyses are now based on different parameter combinations.

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).

Good point.
However, it is hard to do an apple-to-apple comparison. Because we are comparing to raw data, whereas, Kettle et al., 1999 and Lana et al., 2011 interpolated the data, it is hard to extract the raw data information from the climatological map. We thus changed the wording, and weakened the comparison as follow,

“the ability of the ANN to build a nonlinear relationship between DMS and environmental predictors allows it to capture much of the variance”

We also added more comparison to previous model results as shown in Fig. 3. and Fig. 4 in the text, also attached below.

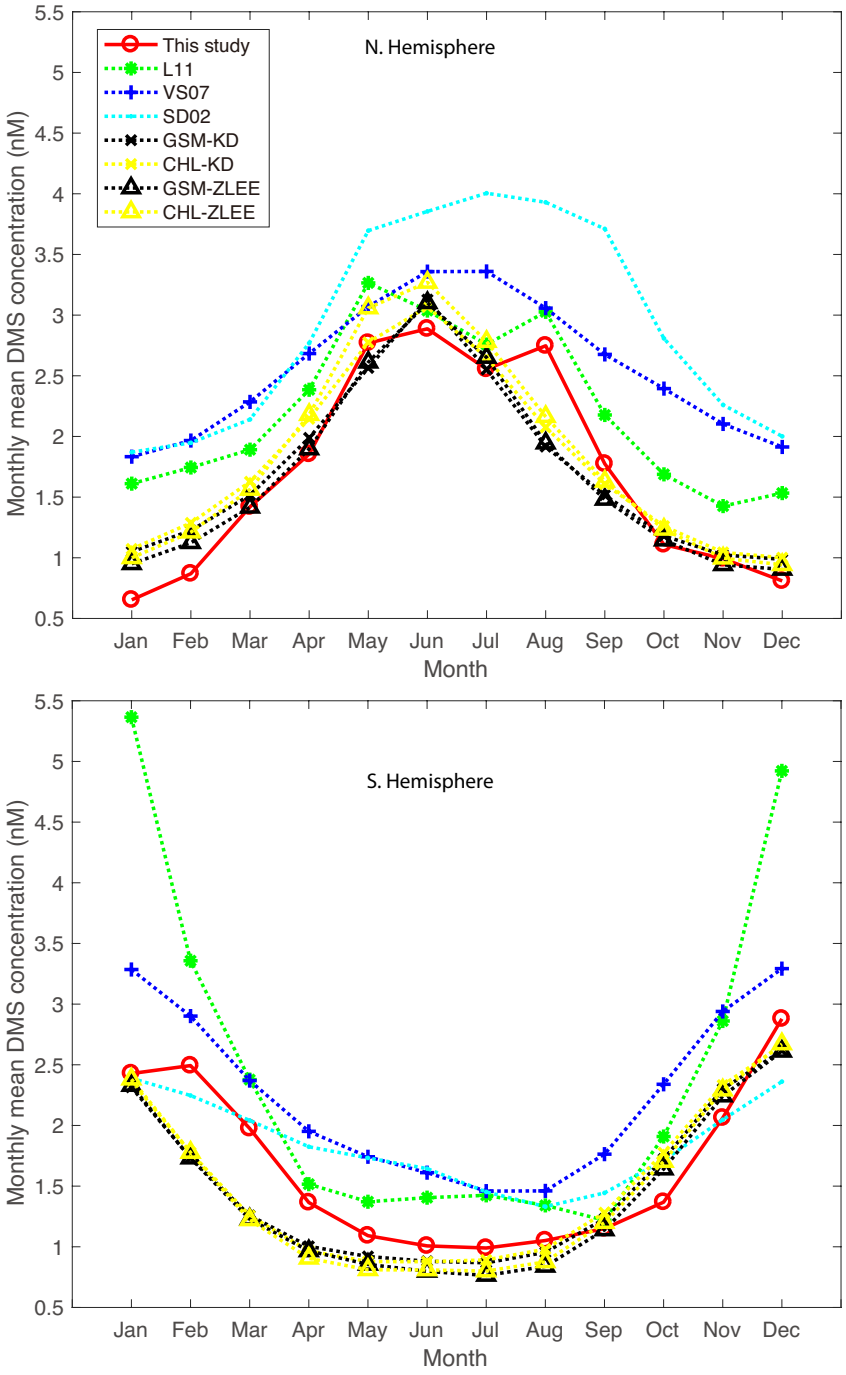

Figure 3. Comparisons of monthly mean DMS concentrations between this study and previous studies (Simó and Dachs, 2002; Vallina and Simó, 2007; Lana et al., 2011; Galí et al., 2018).

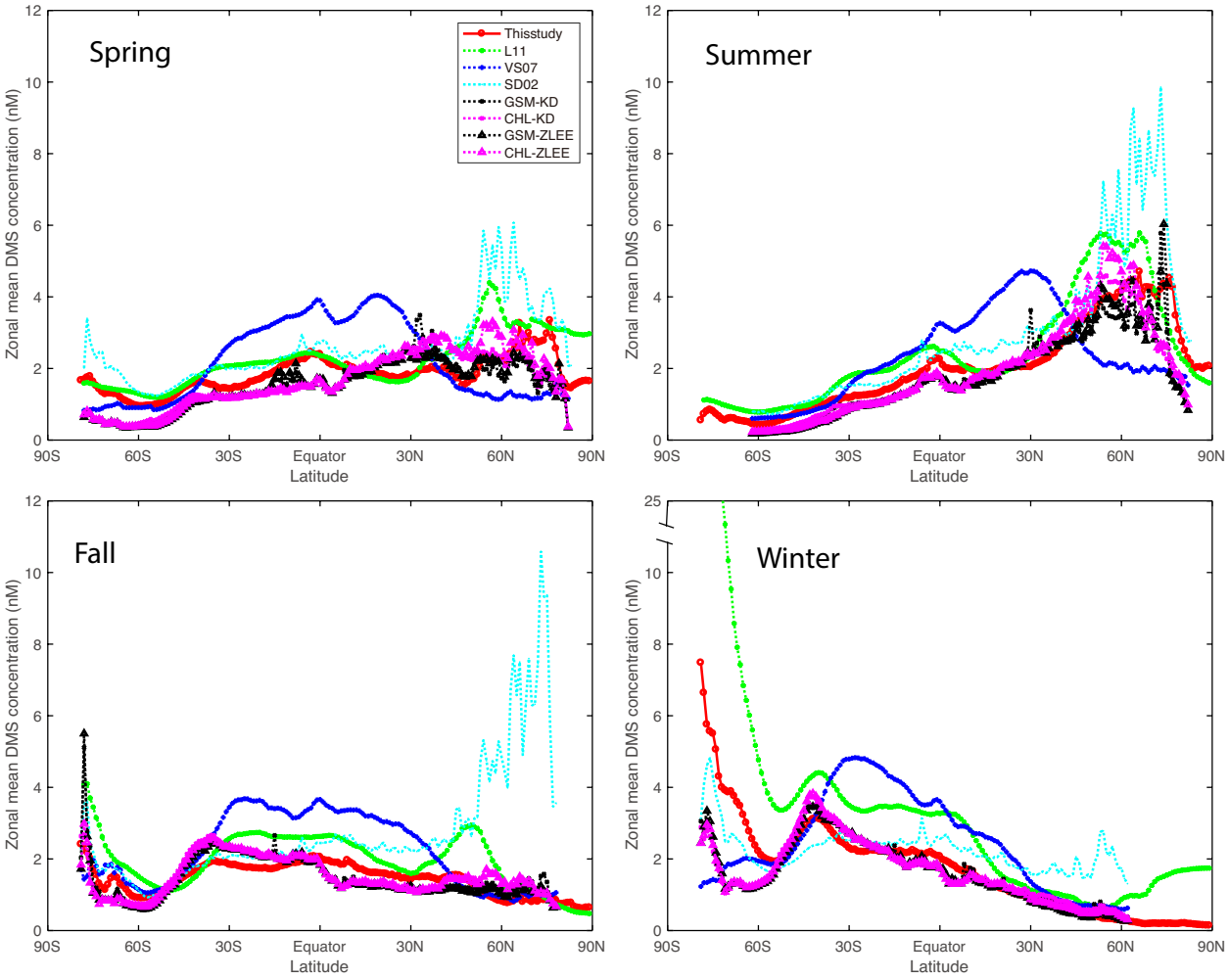

Figure 4. Comparisons of zonally mean DMS concentrations between this study and previous studies (Simó and Dachs, 2002; Vallina and Simó, 2007; Lana et al., 2011; Galí et al., 2018).

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.

The code is previously in a private repository, and now is public (https://github.com/weileiw/ANN-DMS-code). We have also uploaded the corresponding data used to train the model in the following directory: https://zenodo.org/record/3833233#.XsM4cBP0nV4

I also noticed a couple of typos:
p. 2, l. 40: “result” -> “results” or “result[s]”
Corrected, Thank you.
p. 5, l. 31: “deduction” -> “reduction”
Corrected, Thank you.

p. 5, l. 133: “assemble” -> “ensemble” (?)
Corrected, Thank you.
p. 7, l. 189: “wasters” -> “waters”
Corrected, Thank you.

Global ocean dimethyl sulfide climatology estimated from observations and an artificial neural network

Wei-Lei Wang[1], Guisheng Song[2], François Primeau[1], Eric S. Saltzman[1,3], Thomas G. Bell[1,4], and J. Keith Moore[1]

[1]Department of Earth System Science, University of California at Irvine, Irvine, California, USA
[2]School of Marine Science and Technology, Tianjin University, Tianjin, 300072, China
[3]Department of Chemistry, University of California at Irvine, Irvine, California, USA
[4]Plymouth Marine Laboratory, Prospect Place, Plymouth, PL1 3DH, UK

Correspondence: Wei-Lei Wang (weilei.wang@gmail.com)

Abstract. Marine dimethyl sulfide (DMS) is important to climate due to the ability of DMS to alter Earth's radiation budget. Knowledge of the global-scale distribution, seasonal variability, and sea-to-air flux of DMS is needed in order to understand the factors controlling surface ocean DMS emissions. Here we examine the use of an artificial neural network (ANN) to extrapolate available DMS measurements to the global ocean and produce a global climatology with monthly temporal resolution. A global database of 82,996 ship-based DMS measurements in surface waters was used along with a suite of environmental parameters consisting of latitude-longitude coordinates, time-of-day, time-of-year, solar radiation, mixed layer depth, sea surface temperature, salinity, nitrate, phosphate, and silicate. Linear regressions of DMS against the environmental parameters show that on a global scale mixed layer depth and solar radiation are the strongest predictors of DMS. These parameters capture $\sim$9% and $\sim$7% of the raw DMS data variance, respectively. Multi-linear regression can capture more of the raw data variance ($\sim$39%), but strongly underestimates DMS in high concentrations regions. In contrast, the artificial neural network captures $\sim$66% of the raw data variance in our database. Like prior climatologies our results show a strong seasonal cycle in surface ocean DMS with highest concentrations and sea to air fluxes in the high-latitude summertime oceans. We estimate a lower global sea-to-air DMS flux (20.12$\pm$0.43 Tg S yr^{-1}) than the prior estimate based on a map interpolation method (Lana et al., 2011) when the same gas transfer velocity parameterization is used.

1 Introduction

Dimethyl sulfide emitted from the surface ocean is the major precursor for aerosol sulfate in the marine atmosphere. These aerosols play a significant role in the climate system both directly, through aerosol radiative effects and indirectly, through their role as cloud condensation nuclei and influence on cloud radiative properties (Andreae and Rosenfeld, 2008). Assessing the impact of DMS on global climate requires an understanding of the seawater DMS distribution and the factors controlling variability on a variety of spatial and temporal scales. Dimethyl sulfide is produced in surface waters, mainly via enzymatic cleavage of the biogenic compound dimethyl sulfoniopropionate (DMSP; (e.g. Stefels et al., 2007)). The abundance of DMS in surface waters is a function of numerous factors controlling production, loss rates, and pathways of both DMSP and DMS

(Simó, 2001; Toole and Siegel, 2004; Galí et al., 2015). Developing mechanistic and predictive models of surface ocean DMS is challenging due to limitations of the existing observational database and process rate measurements.

Given the biogenic origin of DMS, early efforts focused on the relationship between DMS and Chl a (a proxy for biomass). Positive correlations between DMS and Chl *a* have been reported on basin scales (e.g. Andreae and Barnard, 1984; Yang et al., 1999). However, this positive correlation disappears when more data are used. Kettle et al. (1999) found no significant relationship between DMS and Chl *a* based on the global DMS data set available at the time. The weak relationship may be caused by the so-called "summer DMS paradox", which describes a phenomenon that annual maximum of surface DMS concentration is commonly detected in summer when Chl *a* is at its annual minimum in mid and subtropical low latitude waters (Simó and Pedrós-Alió, 1999). Kettle et al. (1999) also tested linear regression models on a compilation of data, including sea surface salinity and temperature, nitrate, silicate, phosphate, and Chl *a*. The authors then concluded that no simple algorithm based on linear regression could be used to create monthly DMS fields, indicating that more complex mechanisms can control surface DMS concentrations.

Simó and Dachs (2002) achieved a strong linear relationship between heavily binned/averaged DMS and mixed layer depth (MLD) when Chl-*a*/MLD ≥ 0.02, and a logarithmic relationship between DMS and Chl-*a*/MLD when Chl-*a*/MLD < 0.02. Vallina and Simó (2007) found a linear relationship between DMS concentration and solar radiation dose (SRD) in the coastal northwestern Mediterranean. They conducted a global scale study by dividing the ocean into 10° latitude by 20° longitude boxes and correlating SRD and the box averaged DMS concentration. A strong linear relationship was detected in this filtered dataset. Derevianko et al. (2009) reexamined the relationship between SRD/MLD and DMS concentration by using 1° by 1° bins, and found that only a small fraction (14%) of the DMS variance was captured by a linear model based on SRD or MLD. These authors also pointed out that the previously identified strong relationship between MLD/SRD and DMS "results from the reduction in the total variance in the data due to binning" (Derevianko et al., 2009).

Prognostic models have also been used to obtain climatological DMS distributions. In these models, phytoplankton are divided into different groups based on their ability to produce DMS. For example, diatoms produce less DMS than coccolithophores and *Phaeocystis* (e.g. Bopp et al., 2003; Vogt et al., 2010; Gypens et al., 2014). Elliott (2009) implicitly incorporated *Phaeocystis* in a model by assuming that DMS yields are simply related to temperature. The work of Wang et al. (2015) explicitly incorporated *Phaeocystis* into the Biogeochemical Elemental Cycling (BEC) model and included DMSP production from each phytoplankton group, along with DMS leakage pathways from algal cells, (grazing, lysis, and exudation). Despite this level of modeling detail, there are still large discrepancies between the model simulations and in situ measurements (Le Clainche et al., 2010; Tesdal et al., 2016).

The DMS climatologies used in most climate models were obtained by extrapolating observed DMS to the global ocean using objective analysis schemes (Kettle et al., 1999; Lana et al., 2011). In those climatologies, observational data were first binned and averaged into 1° by 1° grid squares, which were then grouped into 57 static biogeographic provinces according to Longhurst (1998). Many provinces lacked adequate data to create a reliable climatology (Fig. A1). In those situations, they first generated an annual cycle with monthly means for each province. Temporal interpolations were used to fill the monthly gaps

if there were enough data to create a robust annual mean. Otherwise, interpolation from neighboring provinces was used to fill the remaining gaps. Major gaps remain in the observational data base for wintertime in the high latitudes of both hemispheres.

Machine learning is being increasingly used in oceanography and geoscience studies (Bergen et al., 2019). For example, Roshan and DeVries (2017) applied an artificial neural network (ANN) to extrapolate observed dissolved organic carbon (DOC) to the global ocean. Rafter et al. (2019) used an ensemble of neural networks to study oceanic δ^{15}N distribution. ANNs have also been used to study DMS on regional scales (e.g. Humphries et al., 2012). The popularity of machine learning partially stems from one of its inherent advantages: it can detect non-linear relationships that traditional linear regression models are unable to capture. The precursor of DMS, DMSP, is mainly produced by marine algae (e.g. Kiene et al., 2000; Curson et al., 2011), and a small fraction of DMSP is transformed to DMS by marine algae and/or bacteria lyases (Simó, 2001; Stefels et al., 2007; Curson et al., 2011; Moran et al., 2012), and mostly as a result of food web interactions (Kiene et al., 2000; Simó, 2001). As a result, we expect that there exists a functional relationship between parameters controlling the growth of phytoplankton or species distribution. The objective of this paper is to explore possible relationships between DMS and environmental variables, with the goal of creating a monthly-resolved DMS climatology.

The paper is organized as follows. We begin by exploring the relationships between DMS concentration and various environmental parameters taken one at a time using linear regression. We then do a stepwise multilinear regression to create a reference model to which we compare our neural network model results. Lastly, we train an ANN using DMS measurements and environmental parameters. With the trained networks, we extrapolate the sparse measurements globally to obtain gridded fields of monthly DMS distributions and sea-to-air DMS fluxes.

2 Materials and Methods

2.1 Data sources and cleaning

Surface ocean DMS data were obtained from the Global Surface Seawater DMS database (PMEL) and from the North Atlantic Aerosol and Marine Ecosystems experiment [NAAMES] (Behrenfeld et al., 2019) (Table A1) . In total, there are 93,571 valid measurements (PMEL: 86,785 and NAAMES: 6,786) after removing ultra-low (<0.1 nM) and ultra-high (>100 nM) DMS measurements according to Galí et al. (2015). The number of measurements used are substantially more than the 47,313 used by Lana et al. (2011). The Global Surface Seawater DMS database also includes some ancillary in situ data, such as DMSP (4,620), Chl *a* (PMEL: 11,491, NAAMES: 6750), sea surface temperature (SST; PMEL: 81,069, NAAMES: 6,786), and salinity (SSS; PMEL: 77,209, NAAMES: 6,786). In situ SST and SSS were used if available. If not, monthly climatology data from other sources (Table A1) are used to fill the gaps. SeaWiFS Chl-*a* data (Level 3-binned, spatial resolution of 9.2 km) from December 1997 to March 2010 were matched to DMS data according to coordinates and sampling date. We compared PMEL in situ Chl *a* to SeaWiFS Chl *a*, which are well correlated on logarithmic scale (R^2 = 0.64) with a slope of 0.67 and an intercept of -0.06, [$log(Chl_{SeaWiFS}) = 0.67 log(Chl_{insitu}) - 0.01$], which means that on logarithmic scale SeaWiFS Chl-*a* concentrations are on average ∼30% lower than those of in situ Chl-*a* concentrations. This is possibly because SeaWiFS Chl *a* is calibrated based on HPLC determined Chl *a* (Morel et al., 2007), which on average is ∼40% lower than that determined

using Fluorometric method (Sathyendranath et al., 2009). Unfortunately, there is no flag in the database showing how Chl *a* was determined. For consistency, we use only Chl-*a* data retrieved from SeaWiFS in the following multilinear and network models.

SeaWiFS photosynthetically available radiation (PAR) and diffuse attenuation coefficient for downwelling irradiance at 490 nm (Kd490) (both are L3BIN with spatial resolution of 9.3 km) from September 1997 to August 2010 were matched with DMS according to coordinates and sampling date. Mixed layer depth climatologies were obtained from the MIMOC climatology (Schmidtko et al., 2013). Sea ice cover was from a simulation with the ocean component of the Community Earth System Model (CESM) forced with a repeating thirty year cycle (1980-2009) of NCEP reanalysis datasets (Wang et al., 2019). The output was averaged into a monthly climatology and was used as part of the air-sea gas exchange calculations. Nutrient data (nitrate, phosphate, and silicate) from World Ocean Atlas (WOA2013, Garcia et al. (2013)) were also included in the multilinear regression and neural network analyses, since they can exert influence on phytoplankton distribution and thus influence DMS production (Wang et al., 2015; Archer et al., 2009). The ancillary data are then matched with DMS data according to sampling location and time of year.

The entire dataset is subjected to another round of quality control following Galí et al. (2015). Specifically, coastal data with salinity lower than 30 and samples with sampling depth greater than 10 m were removed. Additionally, data with extremely low nutrient concentrations (e.g. DIP < 0.01 μM, DIN < 0.01 μM, SiO < 0.1 μM) or low Chl-*a* concentrations (Chl *a* < 0.01 mg/m^3) were also removed because a) the low concentrations are below traditional method detection limits and b) they cause the data distributions severely left skewed, which significantly affects the performance of a ANN model.

2.2 Linear regressions

Linear regression models are conducted on three sets of data to diagnose the predictive skill of each ancillary variable. As a first step, we restrict the regression model to the PMEL data sets where both DMS and the predictor variable are simultaneously available. This selection process yields a total of 10,404 pairs for Chl *a* and DMS, 4,061 pairs of total DMSP (DMSPt) and DMS, 69,197 pairs of SST and DMS, and 85,150 pairs of SSS and DMS, respectively. In a second step, we conduct regression models on combined PMEL and NAAMES data. Since almost all NAAMES samples are accompanied by in situ measurements of Chl *a*, SSS, and SST, the data pairs increased to 17,153 pairs for Chl *a* and DMS, 75,983 pairs of SSS and DMS, and 91,936 pairs of SST and DMS, respectively. In a third step, the unmeasured predictors (i.e. MLD, PAR, Nitrate (DIN), Phosphate (DIP), and Silicate (SiO_4^{4-}), SST, SSS, and Chl *a*) are filled in using monthly climatology data from the previously cited sources. DMSPt is not included, because there is no available climatological dataset to fill the missing values.

To reduce the dynamic range, we log-transform the DMS, DMSPt, Chl *a*, MLD, DIP, DIN, SiO, and SST after conversion to absolute temperature to avoid losing data with temperature below or equal to 0 °C. ~~We do not log-transform SST to avoid losing data with temperature below (equal to) zero.~~ The corresponding predictors are then standardized to their z-score, $Z \equiv (C - \overline{C})/\sigma$, where C is predictor's concentration; $\overline{C}$ is the mean of the variables; and σ is standard deviation of the variables. Matlab's `polyfit` function is applied to each pair to fit a first degree polynomial, i.e. a linear regression.

2.3 Multilinear regression

We begin by applying a step-wise multi-linear regression model to the environmental data using Matlab's `stepwiselm` function. In a first test, we consider a total of eight potential DMS predictors: PAR, MLD, Chl *a*, SSS, SST, DIN, DIP, and SiO. In a second test, we combine the above eight potential parameters with sampling location and time parameters (Eq: 1-3). The ANN requires that the predictor fields be available for every DMS data point so we fill missing values in the environmental dataset with climatological data. We eliminate DMS measurements that are under ice cover, leaving us with 82,996 DMS measurements with a complete set of predictors.

The in situ sampling times (months and hours) were converted to periodic functions using sine and cosine functions to address the data continuity issue, such that in a diurnal or seasonal cycle the start (0th hour or January) and the end (24th hour or December) of a cycle share the same properties, but are numerically different. The coordinate space notations have a similar issue in the longitudinal direction. The conversions are conducted according to Gade (2010) and Gregor et al. (2017) as follows:

$$\begin{bmatrix} \mathrm{H1} \\ \mathrm{H2} \end{bmatrix} = \begin{bmatrix} \cos(\mathrm{hour}\frac{2\pi}{24}) \\ \sin(\mathrm{hour}\frac{2\pi}{24}) \end{bmatrix}, \tag{1}$$

$$\begin{bmatrix} \mathrm{M1} \\ \mathrm{M2} \end{bmatrix} = \begin{bmatrix} \cos(\mathrm{month}\frac{2\pi}{12}) \\ \sin(\mathrm{month}\frac{2\pi}{12}) \end{bmatrix}, \tag{2}$$

$$\begin{bmatrix} \mathrm{L1} \\ \mathrm{L2} \\ \mathrm{L3} \end{bmatrix} = \begin{bmatrix} \sin(\mathrm{lat}\frac{\pi}{180}) \\ \sin(\mathrm{lon}\frac{\pi}{180})\cos(\mathrm{lat}\frac{\pi}{180}) \\ \text{-}\cos(\mathrm{lon}\frac{\pi}{180})\cos(\mathrm{lat}\frac{\pi}{180}) \end{bmatrix}. \tag{3}$$

Bayesian Information Criterion (BIC) of 0.01 is used as a criterion for accepting or rejecting a predictor, which means that predictors are removed if they induce a BIC increase of more than 0.01.

2.4 Artificial Neural Network (ANN)

To assess the possibility that a non-linear model might provide better prediction, we train artificial neural networks (ANNs) using the `Keras` deep learning toolbox in Python. DMS concentration along with the eight environmental predictors (PAR, MLD, Chl *a*, SSS, SST, DIN, DIP, and SiO) are log-transformed. The predictors' dynamic ranges are then constrained to the [-1,1] interval using a minmax normalization, i.e. $C_{norm} \equiv (C - C_{min})/(C_{max} - C_{min})$, where C_{min} and C_{max} are the minimum and maximum values in the data C, respectively.

The dataset is then separated into three sets: training, internal testing, and external validating sets. Data from each of the fourteen one-degree-latitude bands (64°N–65°N, 54°N–55°N, 44°N–45°N, 34°N–35°N, 24°N–25°N, 14°N–15°N, 4°N–5°N, 4°N–5°S, 14°S–15°S, 24°S–25°S, 34°S–35°S, 44°S–45°S, 54°S–55°S, 64°S–65°S,) are left out for inter-

nal testing (9,084 points). Data from each of the fifteen one-degree-latitude bands (69°N−70°N, 59°N−60°N, 49°N−50°N, 39°N−40°N, 29°N−30°N, 19°N−20°N, 9°N−10°N, 1°N−0°S, 9°S−10°S, 19°S−20°S, 29°S−30°S, 39°S−40°S, 49°S−50°S, 59°S−60°S, 69°S−70°S) are left out for external validation (10,870 points). The remaining data (63,042 points) are used to train the neural network. The data was split into sets manually rather than automatically. The online DMS data are organized by contributor ID, and automatic splitting draws a continuous portion from the data. The data portion may come from a specific contributor who collected data from a specific region and it may therefore not be representative. This would result in an over-trained or a under-trained model. To make the selected data more representative, a common practice is to shuffle the data, and then randomly draw a fraction from the shuffled data. For DMS, data collected from the same cruise are highly intercorrelated, so that shuffling and randomly splitting "leaks" information to the model and causes overfitting. We manually adjust the hyper-parameters (dropout ratio, hidden layers, number of nodes etc.) using the data that has been manually-divided into training, internal testing, and external validation subsets. After obtaining a satisfactory combination of those hyper-parameters (as discussed below), we fix them and fine tune the network using all available data.

The network has one input layer with input nodes corresponding to the number of predictors, two dense hidden layers with 128 nodes each, and one output layer with one node corresponding to the predicted logarithm of DMS concentration. To avoid overfitting, we add two dropout layers with a dropout ratio of 25% after each hidden layer. We also apply a L2 kernel regularizer for each hidden layer with the regulation parameter value set to 0.001. When the network is trained, the mean squared error of the internal validation data is monitored, and the training is stopped when there is no error reduction in 10 epochs An epoch consists of one forward pass and one backward pass of all the training examples. Only the best model with the lowest validation mean squared error is saved. We tested different network setups - the current setting achieves goodness of fit, but avoids overfitting.

2.4.1 Parameter selections

The 15 predictors (8 environmental predictors and 7 time and coordination signatures) were tested separately. In the first set of tests, we use only time and location parameters. In the second set of tests, we run a series models that examine every possible combination of the eight environmental parameters (a total of 255 combinations) to time and location parameters. The models are then ranked according to the root mean square error of the validation data.

2.4.2 Monthly climatology

To obtain monthly DMS climatologies, we interpolate the corresponding predictor variables (PAR, MLD, Chl *a*, SSS, SST, DIN, DIP, and SiO) onto a 1° by 1° grid. Coordinates and target months are transformed accordingly. We then apply the top 10 (Section: 2.4.1) trained networks to obtain DMS monthly concentrations. Monthly results from 10 models are then used to produce the final monthly climatology and to analyze uncertainties.

2.5 Sea-to-air flux

Air-sea gas transfer is estimated using the following bulk formula,

$$F = K_w(C_w - C_a/H), \tag{4}$$

where F is sea-to-air gas exchange flux, C_a and C_w are bulk air and bulk water gas concentrations, and K_w (cm/hr) is the overall gas transfer velocity, expressed in water side units (Liss, 1974). K_w reflects the combined resistance to gas transfer on both sides of the interface, as follows:

$$1/K_w = 1/k_w + 1/(Hk_a)), \tag{5}$$

where H is the dimensionless (gas/liquid) Henry's law constant and k_a and k_w are gas transfer velocities in air and seawater. DMS in the surface ocean is strongly supersaturated with respect to that in the overlying atmosphere ($C_w \gg C_a$), which simplifies the flux Eq. 4 to

$$F = K_w C_w, \tag{6}$$

For this study we used two parameterizations for K_w. The Goddijn-Murphy et al. (2012) parameterization (hereafter GM12) is based on regressions between satellite based wind-speed observations with shipboard in situ measurements of DMS gas transfer velocities using eddy covariance. The GM12 parameterization for K_w normalized to a S_c number of 660 is

$$K_{w,660} = 2.1U_{10} - 2.8, \tag{7}$$

where U_{10} is wind speed (m/s) at 10 m above sea surface. We also utilized the Nightingale et al. (2000) (hereafter N00), which is based on shipboard ^{3}He/SF$_6$ dual tracer experiments. Their parameterization for water side only DMS gas transfer velocity at a Schmidt number of 660 ($\kappa_{w,660}$) is calculated as follows,

$$k_{w,660} = (0.222U_{10}^2 + 0.333U_{10})(Sc_{DMS}/600)^{-0.5}, \tag{8}$$

where Sc_{DMS} is calculated as a function of temperature after Saltzman et al. (1993). A total transfer velocity is obtained from N00 as follows,

$$K_{w,660} = k_{w,660}(1 - \gamma_a), \tag{9}$$

where γ_a is atmospheric gradient fraction given by $\gamma_a = 1/(1 + k_a/\alpha k_{w,660})$ (McGillis et al., 2000). Air side DMS transfer velocity is given as $k_a = 659U_{10}(M_{DMS}/M_{H_2O})^{-0.5}$, where M_{DMS} and M_{H_2O} are the molecular weights of DMS and water, respectively (McGillis et al., 2000).

DMS fluxes were calculated using surface ocean DMS concentrations from the ANN results and a satellite-based wind speed climatology (Table A1 and Fig. A2). Because the N00 parameterization was calibrated using in situ wind speeds and has a nonlinear quadratic dependence on wind speed, the use of monthly mean wind speeds will introduce errors. To reconcile the

differences between in situ wind speed and monthly mean wind speed, a correction is applied according to Simó and Dachs (2002) by assuming that instantaneous wind speeds follow a Rayleigh distribution. Eq. 8 thus becomes $k_{w,660} = [0.222\eta^2\Gamma(1+2/\xi) + 0.333\eta\Gamma(s)](Sc_{DMS}/600)^{-0.5}$, where $\eta^2 = 4U_{10}^2/\pi$; $s = (1+1/\xi)$, and ξ = 2 for Rayleigh distribution (Livingstone and Imboden, 1993). Ice fraction data are from the CESM simulation monthly climatology. DMS fluxes from ice-covered regions are set to zero, although DMS concentration in or below sea ice is not necessarily zero.

3 Results and discussion

3.1 Linear regressions

The linear regression coefficients and R^2 values are summarized in Table 1. For the test using in situ measurements, DMS and DMSPt show the strongest positive correlation with a R^2 value of 0.41 (n = 4061). Galí et al. (2018) reported a slightly higher R^2 value (0.42) with less data points (n = 3637). It is not surprising to find the strong relationship between total DMSP (DMSPt) and DMS, since DMS derives from the enzymatic cleavage of DMSP (Stefels, 2000; Stefels et al., 2007). ~~However, it is difficult to infer global scale climatological DMS distributions from total DMSP measurements, because there are approximately 10-fold fewer measurements of total DMSP compared to DMS.~~ Since DMSP is directly produced by phytoplankton and does not undergo sea-to-air gas exchange, it is relatively easy to parameterize in a biogeochemical model (Galí et al., 2015). The strong relationship between DMS and DMSP point toward a potential way to model marine seawater DMS. McParland and Levine (2019) developed a mechanistic model that related intracellular DMSP concentration to environmental stress, and coupled the model with MIT ecosystem model (DARWIN) to estimate global ocean DMSP distribution. Galí et al. (2015) first applied a remote sensing algorithm to obtain a DMSP climatology, from which they predict DMS climatology through an empirical relationship with PAR (Galí et al., 2018).

The second strongest predictor is in situ Chl *a* (R^2 = 0.21, n = 10,404), which is slightly higher than that by Galí et al. (2018) who reported a R^2 value of 0.20 (n = 8,141). The positive correlation between Chl *a* and DMS is possibly due to the fact that the precursor of DMS, namely DMSP, is biogenic. However, when we test the relationship on satellite-based climatological Chl *a*, it becomes weaker (PMEL, R^2 = 0.09, n = 81,767; PMEL+NAAMES R^2 = 0.09, n = 88,516). The weaker relationship can be caused by several reasons: 1) Greater variance in the larger dataset (81,767 vs 10,404); 2) mismatch between satellite derived Chl-*a* concentrations and analytical Chl-*a* concentrations; 3) the in situ Chl-*a* samples in PMEL database were collected mainly in highly productive regions (Galí et al., 2018), whereas the relationship between Chl-*a* and DMS may negatively correlated in oligotrophic oceans over the seasonal cycle (Galí and Simó, 2015).

When tested against climatological data with gaps filled-in, PAR has the strongest correlation with DMS (PMEL: R^2 = 0.07, n = 82,137; PMEL+NAAMES: R^2 = 0.09, n = 88,923) with a positive correlation slope. Climatological MLD is the second strongest predictor (PMEL: R^2 = 0.06, n = 81,646; PMEL+NAAMES: R^2 = 0.07, n = 88,214) of raw DMS data, with a slope of -0.25 for PMEL and -0.26 for PMEL and NAAMES combined data.

3.2 Multilinear regression

A multilinear regression model that uses a combination of predictors or product of predictors has higher predictive ability than a linear regression model. For example, a multilinear regression model using eight environmental parameters has a R^2 value of 0.28, which is higher than that of any of the linear models. By adding time and location parameters, the R^2 value increases to 0.39 (n = 82,996, Fig. 1.) The results emphasize the importance of including time and location information in the model. Sampling time and location are useful predictors, especially when the output has strong seasonality such as DMS. Given a location and sampling time, the model roughly predicts the level of DMS concentrations (e.g. high latitude DMS concentrations are higher in summer than in winter). However, it is apparent that the multilinear regression model significantly underestimates high DMS concentrations. The generally low correlation coefficient hinders the possibility of reliably extrapolating the model to the global ocean.

3.3 ANN

Fig. 1b displays the ~~tracer-tracer~~ correlation between DMS observations and ANN predictions. Compared to simple linear and multilinear regression models, ANN captures much more of the observed DMS variance (R^2 = 0.66, n= 82,996). Compared to previous extrapolations (Kettle et al., 1999; Lana et al., 2011), the ability of the ANN to build a nonlinear relationship between DMS and environmental predictors allows it to capture ~~more~~ much of the variance. It also incorporates diurnal and seasonal signals present in the data. As a result, the extrapolation obtained from the ANN considers the relationship with geographical neighbors and also with temporal relationships.

From traditional linear or multilinear models, one can easily determine which parameter is a strong predictor and how a predictor influences the state variable (e.g. the correlation between DMSP and DMS). An ANN model is much more complex, it adjusts weights of each node that connect inputs and outputs. The relationship between inputs and outputs is therefore much more subtle and is why ANN models are generally referred to as a "Black Box". In this study, we design experiments that help open this "Black Box" and reveal parameters that drive surface ocean DMS distributions.

As shown in Fig. 2, without using any environmental parameters, sampling location and date alone can explain 44% of the validation data variance (RMSE = 0.65 on natural logarithm scale). Given the strong correlation between solar radiation and DMS concentration reported by Vallina and Simó (2007), one would expect that adding sampling time would improve the model performance. However, it increases RMSE slightly (Fig. 2a). Galí et al. (2013c) studied diel cycle at the Mediterranean Sea and Sargasso Sea. Among their four experiments (three in the Mediterranean Sea and one in the Sargasso Sea) regular diel variation was observed at only one experiment in the Mediterranean Sea at summer season, with highest DMS values observed at midnight and lowest values at midday. In all the other experiments, diel variations for both DMS and DMSPt pools were small. Gross community DMS production during the daytime was two to three times higher than that in the nighttime, but the high DMS production was compensated by greater photochemical and microbial consumption (Galí et al., 2013c). The balance between DMS production and consumption appears to dampens DMS diel variation. This may explain why adding time parameters does not improve the ANN model's predictive ability.

Adding environmental parameters can further improve the model performance, however, different parameter combinations show different predictive abilities. Among the top 10 models ranked according to RMSE of validation data (PAR + MLD + SAL + SST, MLD + SST, SAL + SST + DIP + Chl *a*, MLD + SST + DIP, PAR + MLD + SAL + SST + SiO + DIP, PAR + MLD + SST + SiO, MLD + SAL + DIP, PAR + MLD + SST + Chl *a*, PAR + MLD + SST + SiO + DIP, SAL + SST + SiO + Chl *a*), 9 models have SST, 8 models have MLD, 5 models have PAR, SSS, and DIP, 4 models have SiO, and 3 models have Chl *a* as a predictor, and none of the models have DIN as a predictor.

Based on the appearance frequency, SST (9 times over top 10 models) should be a strong predictor. Physiologically, DMSP is an important cryoprotectant that helps algae deal with cold temperature especially in high latitude oceans (Thomas and Dieckmann, 2002). However, the linear regression models show that there is almost no correlation between SST and DMS. Also, SST alone with date and location parameters have very low prediction ability (ranked 244 over 255 models). When combined with other parameters, SST helps to improve the model performance. For example, the combination of SST and MLD ranks 2nd place among all models. Therefore, SST may work synergistically with other parameters to increase the prediction ability.

MLD is another important predictor. High DMS concentrations in the open ocean have been detected when the water column is most stratified (Simó and Pedrós-Alió, 1999). The authors proposed that a stratified (high light) environment nourishes strong DMSP producers, or that phytoplankton cellular DMSP quota increases in such an environment. High conversion rates from DMSP to DMS in stratified waters is another reason for high DMS concentrations when MLD is shallow. Meanwhile, the biological DMS consumption rate decreases in oligotrophic oceans (Galí and Simó, 2015). A dilution model was also proposed to explain the anti-correlation between DMS concentration and MLD (Aranami and Tsunogai, 2004). The authors proposed that mixed layer deepening entrains water with little or no DMS into surface waters and dilutes surface DMS concentrations, but recent studies have shown that DMS loss rate via vertical mixing is orders of magnitude lower than production/consumption rates (e.g. Galí et al., 2013c; Royer et al., 2016).

Physiologically, the correlation between PAR and DMS can be explained by two reasons. First, high radiation negatively influences the bacterial population/activity, which decreases DMS consumption (Galí et al., 2013a, b, c; Royer et al., 2016). Second, high radiation promotes DMS production by inducing oxidative stress within algal cells (Toole et al., 2006; Sunda et al., 2002; Royer et al., 2016). Strong correlation between monthly binned and averaged solar radiation dose (SRD) and DMS concentration has been reported (R^2 = 0.94) at the Blanes Bay Microbial Observatory located in the coast of northwest Mediterranean (Vallina and Simó, 2007). Galí et al. (2018) also combined PAR with DMSPt distribution to predict DMS climatologies, indicating that PAR is an important parameter controlling the conversion of DMSPt to DMS.

The appearance of SSS in the models may be due to osmoregulation function of DMSP in cells, which helps algae survive in high salinity waters (Thomas and Dieckmann, 2002; Webb et al., 2019). The appearance of DIP and SiO in the model is probably related to nutrient stress, which can increase DMSP production by low DMSP producers (e.g. diatoms) (McParland and Levine, 2019). In their recent paper, McParland and Levine (2019) showed that intracellular DMSP concentrations in low DMSP producers can increase by ~16 times in nutrient stress situations, while in high DMSP producers (e.g. coccolithophores)

the change is small (1.5-fold). However, they also pointed out that the intracellular changes of DMSP due to nutrient stress has a minor effect on large scale DMSP distribution, and that community composition plays the most important role.

Previous studies of the relationship between DMS and Chl *a* have produced contradictory results. Our linear regression models show a higher correlation coefficient using the in situ data than when using the satellite Chl-*a* data. Strong correlation relationships have been reported in basin scale studies (e.g. Yang et al., 1999). On the other hand, there are numerous studies that observed no correlation between DMS and Chl *a* (e.g. Dacey et al., 1998; Kettle et al., 1999; Toole and Siegel, 2004). The inconsistent relationships indicate the complexity of the biogeochemical reduced sulfur cycle. As suggested by Simó (2001), not only can phytoplankton biomass, taxonomy, and activity influence DMS production, but so does food-web structure and dynamics. The inconsistent relationship may also explain the low ranking of Chl *a* in the models.

3.4 Binned data versus raw data

Simó and Dachs (2002) obtained high R^2 values between DMS concentration and the ratio of Chl *a* to MLD (Chl/MLD) when Chl/MLD is greater than or equal to 0.02, and between DMS concentration and ln(MLD) when Chl/MLD is less than 0.02. We tried exactly the same model on raw PMEL data with in situ Chl-*a* measurements and climatological MLD, and found that both correlations between DMS and Chl/MLD (n = 4,921, $R^2 =\sim 0$ .1) and between DMS and ln(MLD) (n = 5,978, $R^2 =\sim 0$) are statistically insignificant. To reduce interannual variability, we binned in situ Chl *a* and DMS into monthly $1^\circ \times 1^\circ$ grid, and retested the above model on the binned data, and found that the correlations are still statistically insignificant.

Vallina and Simó (2007) reported an R^2 of 0.95 (n=14) between DMS concentration and SRD. We applied the same linear regressions on both raw data and monthly $1^\circ \times 1^\circ$ data, and found no significant correlations between DMS and SRD as calculated according to Vallina and Simó (2007):

$$\mathrm{SRD} = \mathrm{SI} \cdot \frac{1}{\mathrm{Kd490} \cdot \mathrm{MLD}} (1 - e^{-\mathrm{Kd490} \cdot \mathrm{MLD}}), \tag{10}$$

where SI is shortwave irradiance (W m^{-2}), which is converted from PAR according to Galí and Simó (2015).

Compared to Simó and Dachs (2002) and Vallina and Simó (2007), we used significantly more data points. For example, in this study, there is a total of 10,899 DMS measurements accompanied with simultaneous Chl *a* measurements versus 2,385 data points used in Simó and Dachs (2002), and 83,152 (DMS, MLD) pairs in this study versus 26,400 in Vallina and Simó (2007). Another noticeable difference between the current study and previous analyses is that both Simó and Dachs (2002) and Vallina and Simó (2007) binned the data into large longitude and latitude grids. By doing so, the raw data variance is greatly reduced.

Binning data will necessarily result in loss of information. A lot of information is associated with sampling location and date as shown in Fig. 2a. By binning the data into monthly $1^\circ \times 1^\circ$ grid, the number of data points decreases from 82,996 to only 9,018; sampling date feature (365) will be average to 12 months, and coordination combinations will be averaged from 87,332x87,332 to 180°x360°, which represents a substantial loss of information. For ANN models, using less data points can lead to overfitting (See Fig. 2b).

3.5 DMS distributions

Northern and Southern hemisphere monthly mean DMS concentrations are plotted along with results from previous studies (Simó and Dachs, 2002; Vallina and Simó, 2007; Lana et al., 2011; Galí et al., 2018) (Fig. 3a). Overall, all models show similar seasonal patterns with highest concentrations in summer and lowest concentrations in winter. Our predictions are highly consistent with the products derived from satellite data reported by Galí et al. (2018), who used an optimized relationship between DMS, DMSPt, and PAR to obtain DMS climatology from satellite retrieved PAR and DMSPt fields (Galí et al., 2015). In the northern hemisphere, the algorithms by Simó and Dachs (2002) (SD02 hereafter) and by Vallina and Simó (2007) (VS07 hereafter) generate higher and flatter seasonality. From zonal average plots (Fig.4), it is clear that the elevated monthly means from SD02 are caused by high concentrations in high latitude oceans, whereas, high monthly means of VS07 are caused by high DMS concentrations in low and middle latitude. High DMS concentration in high latitude summer (SD02) is driven by shoaling of MLD caused by freshwater runoff (Galí et al., 2018), while high DMS concentrations at low/middle latitude (VS07) are driven by strong solar radiation dose, which is a joint effect of shallow MLD and strong irradiance.

L11 stands out in the S. hemisphere monthly mean plot (Fig. 3b), with the highest mean concentrations in January and December, when DMS concentrations are $\sim$2 times higher than other model predictions. Galí et al. (2018) identified five shortcomings associated with the direct interpolation method employed by Lana et al. (2011). All shortcomings concern the nature of in situ DMS data, including right-skewed distribution, lack of spatial and temporal coverage, lack of duplicate measurements, and sampling bias towards DMS-productive conditions. Because of the sparsity and skewed distribution, the interpolation/extrapolation method broadcasts small scale features to large scales (Tesdal et al., 2016). This is especially true for the month of January and December when the elevated L11 monthly means were mainly driven by a small amount of extremely high DMS measurements (>40 nM) near the Antarctic continent. On the other hand, empirical models including the ANN model used in this study rely on environmental parameter climatologies to obtain the DMS climatology. Extreme conditions are smoothed out in climatological data, e.g. in the DMS database the maximum in situ Chl-*a* concentration is > 800 mg/m^3, whereas it is $\sim$50 mg/m^3 in the SeaWiFS climatology. When climatological data are used to generate DMS distribution, a smaller variance than in situ data is expected.

Fig. 5 displays monthly DMS concentration distributions predicted by the ANN. Generally, DMS concentrations in polar regions show strong seasonality. The highest DMS concentrations are in summer when light and temperature are ideal for primary production. For example, in austral summer, the Southern Ocean circumpolar regions display the highest DMS concentration ($>$10 nM). DMS concentration in the Scotia Sea and Ross Sea display the highest DMS concentration, which gradually decreases and falls below 0.5 nM in the following months when primary production is limited by light or low temperature. In boreal summer, DMS concentration in the Bering Sea and Greenland Sea can exceed 20 nM.

The summertime high DMS concentration at high latitudes is believed to be linked to the release of ice algae that are prolific DMSP producers (Stefels et al., 2012; Webb et al., 2019). As an important cryoprotectant and osmolyte, DMSP helps ice algae to cope with the low temperature and high salinity conditions (Thomas and Dieckmann, 2002). High DMS concentrations at high latitudes have also been observed to accompany blooms of coccolithophores and *Phaeocystis*, which are strong DMSP

producers (Neukermans et al., 2018; Wang et al., 2015). The shoaling of mixed layer depth in summer provides favorable conditions, i.e. stable and warm, with adequate irradiation for coccolithophores and *Phaeocystis* growth (Galí et al., 2019).

Another interesting region is the Pacific equatorial upwelling region. Large-scale upwelling brings nutrient-rich waters to the surface, which nourish highly productive phytoplankton communities. Overall, the seasonality in the equatorial Pacific is weaker than that in polar regions, but there is still a clear seasonal pattern. In the period from December to April, the tongue with higher DMS concentration ($\sim$3 nM) extends to the west Pacific Ocean reaching the east coast of Australia and the Philippine Sea. The tongue gradually retreats eastward in the following months. From September to November, the tongue is constrained to the eastern Pacific and DMS concentration falls to its lowest values ($<$2.0 nM). High DMS concentrations in the west Pacific ocean from November to February are also predicted by Lana et al. (2011).

The subtropical gyres show consistently low DMS concentrations and weak seasonal cycles throughout the year. In the southern hemisphere gyres, DMS concentrations are highest during austral summer, when the ocean is strongly stratified and local primary production is low. There are hot spots where DMS concentration exceeds 3 nM in December and February. DMS concentrations are generally low ($\leq$ 1 nM) during austral spring and winter seasons. In the period from April to September, DMS concentrations in the S. Atlantic gyre fall below ~~0.1 nM~~ 0.6 nM. In the northern hemisphere gyres, DMS concentrations are high during the boreal summer season. Fig. 6 compares monthly mean Chl-*a* concentrations to DMS concentrations in N. and S. hemisphere gyres. The concentrations are normalized to the range of 0 to 1. It is clear that Chl *a* and DMS are anti-correlated, DMS concentration peaks at summer season when Ch-*a* concentration is generally low. This phenomenon is previously termed as "summer DMS paradox" (Simó and Pedrós-Alió, 1999). This pattern is more apparent in the S. hemisphere gyres, because the terrestrial influence is smaller in the S. hemisphere than in the N. hemisphere.

3.6 Sea-to-air flux

In this study, we computed monthly sea-to-air DMS fluxes using both the GM12 and N00 gas transfer velocity parameterizations (Fig. 7 and Fig. 8). These yield global DMS annual fluxes of 15.89$\pm$0.34 Tg S yr^{-1} [GM12] and 20.12$\pm$0.43 Tg S yr^{-1} [N00], respectively. The uncertainties ($\pm 1\sigma$) are calculated according to DMS distributions from the top 10 ANN models based on different parameter combinations. We also calculated sea-to-air DMS fluxes using the N00 parameterization and previous DMS climatologies from Lana et al. (2011) [L11], Simó and Dachs (2002) [SD02], Vallina and Simó (2007) [VS07], and four from Galí et al. (2018)[Gali18]. Among those climatologies, VS07 produces the highest annual DMS flux (31.59 Tg S yr^{-1}), the ensemble of Galí et al. (2018) climatologies produce the lowest flux (18.18 $\pm$ 0.52 Tg S yr^{-1}) (Table 2). Generally, our fluxes are consistent with previous results when the same flux parameterization, wind speed field, sea surface temperature, and ice coverage are used. The sea-to-air flux based on the GM12 parameterization is $\sim$24% lower than that based on N00.

Geographically, in the high-latitude northern hemisphere, sea-to-air DMS fluxes are low in boreal winter, even though wind speeds are high. The DMS flux tends to increase in the proceeding months and reaches a maximum in boreal summer, despite the lower wind speeds (Fig. A2). The inverse relationship between wind speed and DMS flux indicates that the high DMS flux is mainly driven by high seawater DMS concentrations. Large sea-to-air DMS fluxes at high latitudes in austral summer are driven jointly by high DMS concentrations and high wind speeds (Fig. 7 and Fig. A2). The eastern tropical Pacific Ocean

displays a year-round intermediate sea-to-air DMS flux. This is mainly driven by the high DMS concentration in this region, since the wind speeds here are generally low (Fig. 7 and Fig. A2).

Fig. 8 displays integrated monthly global DMS fluxes for both hemispheres and for the global ocean based on GM12 velocity parameterizations. Globally, DMS fluxes are highest in the winter months (Dec., Jan., and Feb.) and March, which is mainly driven by high DMS flux in the Southern Hemisphere. There is another peak in the months of July and August because of northern hemisphere flux peaks. An interesting feature is that the Northern hemisphere peak is close to Southern hemisphere though, and does not reach the peak level in the Southern hemisphere. This is mainly because of the larger surface area in the Southern hemisphere. High DMS fluxes in the southern hemisphere have profound impact to the Earth's climate because there are less terrestrial and anthropogenic aerosol inputs compared to the northern hemisphere.

4 Conclusions

The artificial neural network (ANN) used in this study has some advantages compared to the prior methods used to develop DMS climatologies. Most importantly, the ANN utilizes available measurements to fill regions without DMS observations, using non-linear relationships trained in more data rich regions/seasons. By contrast, objective interpolation methods are spatial/temporal averages of sparse data with ~~no~~ weak underlying basis in environmental variability. As a result, the ANN approach captures significantly more of the raw data variance than simple linear/multilinear models. Simple models achieve comparable fits only after heavily binning the DMS observations (e.g. Simó and Dachs, 2002; Galí et al., 2015; Vallina and Simó, 2007; Galí et al., 2018). The ANN is computationally more expensive than the linear/multilinear models, but considerably less expensive than prognostic biogeochemical models (e.g. Vogt et al., 2010; Wang and Moore, 2011; Wang et al., 2015). The principle weakness of the ANN approach is that it does not easily provide scientific insight into the relationships between the parameters. Some insight could be gained by running sensitivity tests in which the response to perturbation of a single parameter is diagnosed.

The ANN approach is a useful tool for developing trace gas climatologies. It may also be useful as a means of assessing the sensitivity of DMS to past/future changes in climate by coupling the ANN to prognostic biogeochemical models. Caution is warranted in the interpretation of such efforts because there is as yet no basis for assessing whether the relationships obtained by training on contemporary measurements apply to the past or will hold in the future. Such relationships could be investigated using paleoceanographic and ice core data (Osman et al., 2019).

The annual sea-to-air DMS flux calculated in this study is slightly ($\sim$23%) lower than the objective interpolation method of Lana et al. (2011) using the same sea-to-air gas exchange models. DMS concentrations from this study are similar to Lana et al. (2011) where measurements are abundant, so we infer that the difference is likely caused by positive bias in the objective interpolation method for data-sparse regions/seasons.

Code availability. Code for ANN model is available at: https://github.com/weileiw/ANN-DMS-code

Data availability. The data for DMS concentrations and sea-to-air flux are available at DOI: 10.5281/zenodo.3631875.

Author contributions. W.L.W and G.S. initiated the study and drafted the manuscript. W.L.W. built the model with inputs from F.P., E.S.S., and J.K.M.. E.S.S and T.G.B provided new N. Atlantic DMS measurement data. All authors contributed to review the manuscript, and to interpret the data presented.

Competing interests. The authors declare that they have no competing financial interests.

Acknowledgements. We thank the observational DMS community for making their measurements publicly available. J.K.M., F.P., and W.L.W. are supported by DOE Earth System Modeling program (DE-SC0016539). G.S is supported by the Natural Key Research and Development Program of China (2017YFC1404403). E.S.S. and T.G.B. are supported by the NASA North Atlantic Aerosols and Marine Ecosystems Study (NAAMES), which was funded through the NASA Earth Venture Suborbital program. (NNX#15AF31G). The sources of ancillary data are listed in Table A1.

References

Andreae, M. and Rosenfeld, D.: Aerosol–cloud–precipitation interactions. Part 1. The nature and sources of cloud-active aerosols, Earth-Sci Rev, 89, 13–41, 2008.

Andreae, M. O. and Barnard, W. R.: The marine chemistry of dimethylsulfide, Mar. Chem., 14, 267–279, https://doi.org/10.1016/0304-4203(84)90047-1, 1984.

Aranami, K. and Tsunogai, S.: 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, 2004.

Archer, S. D., Cummings, D. G., Llewellyn, C. A., and Fishwick, J. R.: Phytoplankton taxa, irradiance and nutrient availability determine the seasonal cycle of DMSP in temperate shelf seas, Mar Ecol Prog Ser, 394, 111–124, 2009.

Behrenfeld, M. J., Moore, R. H., Hostetler, C. A., Graff, J., Gaube, P., Russell, L. M., Chen, G., Doney, S. C., Giovannoni, S., Liu, H., et al.: The North Atlantic Aerosol and Marine Ecosystem Study (NAAMES): Science Motive and Mission Overview, Front Mar Sci, 6, 2019.

Bergen, K. J., Johnson, P. A., De Hoop, M. V., and Beroza, G. C.: Machine learning for data-driven discovery in solid Earth geoscience, Science, 363, https://doi.org/10.1126/science.aau0323, 2019.

Bopp, L., Aumont, O., Belviso, S., and Monfray, P.: Potential impact of climate change on marine dimethyl sulfide emissions, Tellus, Series B: Chemical and Physical Meteorology, 55, 11–22, https://doi.org/10.1034/j.1600-0889.2003.042.x, 2003.

Curson, A. R., Todd, J. D., Sullivan, M. J., and Johnston, A. W.: Catabolism of dimethylsulphoniopropionate: microorganisms, enzymes and genes, Nature Reviews Microbiology, 9, 849–859, 2011.

Dacey, J. W., Howse, F. A., Michaels, A. F., and Wakeham, S. G.: Temporal variability of dimethylsulfide and dimethylsulfoniopropionate in the Sargasso Sea, Deep Sea Research Part I: Oceanographic Research Papers, 45, 2085–2104, 1998.

Derevianko, G. J., Deutsch, C., and Hall, A.: On the relationship between ocean DMS and solar radiation, Geophys Res Lett, 36, 2–5, https://doi.org/10.1029/2009GL039412, 2009.

Elliott, S.: Dependence of DMS global sea-air flux distribution on transfer velocity and concentration field type, J. Geophys. Res., 114, G02 001, https://doi.org/10.1029/2008JG000710, 2009.

Frouin, R., McPherson, J., Ueyoshi, K., and Franz, B. A.: A time series of photosynthetically available radiation at the ocean surface from SeaWiFS and MODIS data, in: Remote Sensing of the Marine Environment II, vol. 8525, p. 852519, International Society for Optics and Photonics, 2012.

Gade, K.: A Non-singular Horizontal Position Representation, Journal of Navigation, 63, 395–417, https://academic.microsoft.com/paper/2141730410, 2010.

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

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

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

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

Galí, M., Devred, E., Levasseur, M., Royer, S.-J., and Babin, M.: A remote sensing algorithm for planktonic dimethylsulfoniopropionate (DMSP) and an analysis of global patterns, REMOTE SENS ENVIRON, 171, 171–184, 2015.

Galí, M., Levasseur, M., Devred, E., Simó, R., and Babin, M.: Sea-surface dimethylsulfide (DMS) concentration from satellite data at global and regional scales, Biogeosciences, 15, 3497–3519, https://doi.org/10.5194/bg-15-3497-2018, 2018.

Galí, M., Devred, E., Babin, M., and Levasseur, M.: Decadal increase in Arctic dimethylsulfide emission, Proceedings of the National Academy of Sciences, 116, 19 311–19 317, 2019.

Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O. K., Zweng, M. M., Reagan, J. R., and Johnson, D. R.: World Ocean Atlas 2013, Volume 4 : Dissolved Inorganic Nutrients (phosphate, nitrate, silicate), NOAA Atlas NESDIS 76, 4, 25, https://doi.org/10.1182/blood-2011-06-357442, 2013.

Goddijn-Murphy, L., Woolf, D. K., and Marandino, C.: Space-based retrievals of air-sea gas transfer velocities using altimeters: Calibration for dimethyl sulfide, J GEOPHYS RES: OCEANS, 117, 2012.

Gregor, L., Kok, S., and Monteiro, P. M.: Empirical methods for the estimation of Southern Ocean CO2: Support vector and random forest regression, Biogeosciences, 14, 5551–5569, https://doi.org/10.5194/bg-14-5551-2017, 2017.

Gypens, N., Borges, A. V., Speeckaert, G., and Lancelot, C.: The dimethylsulfide cycle in the eutrophied Southern North Sea: A model study integrating phytoplankton and bacterial processes, PLoS ONE, 9, https://doi.org/10.1371/journal.pone.0085862, 2014.

Humphries, G. R., Deal, C. J., Elliott, S., and Huettmann, F.: Spatial predictions of sea surface dimethylsulfide concentrations in the high arctic, Biogeochemistry, 110, 287–301, 2012.

Kettle, A. J., Andreae, M. O., Amouroux, D., Andreae, T. W., Bates, T. S., Berresheim, H., Bingemer, H., Boniforti, R., Curran, M. A., DiTullio, G. R., Helas, G., Jones, G. B., Keller, M. D., Kiene, R. P., Leek, C., Levasseur, M., Malin, G., Maspero, M., Matrai, P., McTaggart, A. R., Mihalopoulos, N., Nguyen, B. C., Novo, A., Putaud, J. P., Rapsomanikis, S., Roberts, G., Schebeske, G., Sharma, S., Simó, R., Staubes, R., Turner, S., and Uher, G.: A global database of sea surface dimethylsulfide (DMS) measurements and a procedure to predict sea surface DMS as a function of latitude, longitude, and month, Global Biogeochem. Cycles, 13, 399–444, https://doi.org/10.1029/1999GB900004, 1999.

Kiene, R. P., Linn, L. J., and Bruton, J. A.: New and important roles for DMSP in marine microbial communities, J. Sea Res., 43, 209–224, 2000.

Lana, A., Bell, T. G., Simó, R., Vallina, S. M., Ballabrera-Poy, J., Kettle, A. J., Dachs, J., Bopp, L., Saltzman, E. S., Stefels, J., Johnson, J. E., and Liss, P. S.: An updated climatology of surface dimethlysulfide concentrations and emission fluxes in the global ocean, Global Biogeochem. Cycles, 25, https://doi.org/10.1029/2010GB003850, 2011.

Le Clainche, Y., Vézina, A., Levasseur, M., Cropp, R. A., Gunson, J. R., Vallina, S. M., Vogt, M., Lancelot, C., Allen, J. I., Archer, S. D., et al.: A first appraisal of prognostic ocean DMS models and prospects for their use in climate models, Global biogeochemical cycles, 24, 2010.

Liss, P. S.: Flux of gases across the air-sea interface, Nature, 247, 181–184, 1974.

Livingstone, D. M. and Imboden, D. M.: The non-linear influence of wind-speed variability on gas transfer in lakes, Tellus B: Chemical and Physical Meteorology, 45, 275–295, 1993.

Longhurst, A. R.: Ecological geography of the sea, Elsevier, 1998.

McGillis, W., Dacey, J., Frew, N., Bock, E., and Nelson, R.: Water-air flux of dimethylsulfide, J GEOPHYS RES: OCEANS, 105, 1187–1193, 2000.

McParland, E. L. and Levine, N. M.: The role of differential DMSP production and community composition in predicting variability of global surface DMSP concentrations, Limnology and Oceanography, 64, 757–773, 2019.

Moran, M. A., Reisch, C. R., Kiene, R. P., and Whitman, W. B.: Genomic insights into bacterial DMSP transformations, Annual review of marine science, 4, 523–542, 2012.

Morel, A., Huot, Y., Gentili, B., Werdell, P. J., Hooker, S. B., and Franz, B. A.: 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, 69–88, 2007.

NASA: SeaWinds on QuickSCAT Level 3 surface wind speed for climate model comparison, Ver. 1. PO.DAAC, CA, USA, https://doi.org/https://doi.org/10.5067/QSSWS-CMIP1, 2012.

NASA: Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group, Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Chlorophyll Data, https://doi.org/data/10.5067/ORBVIEW-2/SEAWIFS/L3M/CHL/2018, 2018.

Neukermans, G., Oziel, L., and Babin, M.: Increased intrusion of warming Atlantic water leads to rapid expansion of temperate phytoplankton in the Arctic, Global change biology, 24, 2545–2553, 2018.

Nightingale, P. D., Malin, G., Law, C. S., Watson, A. J., Liss, P. S., Liddicoat, M. I., Boutin, J., and Upstill-Goddard, R. C.: In situ evaluation of air-sea gas exchange parameterizations using novel conservative and volatile tracers, Global Biogeochem. Cycles, 14, 373–387, 2000.

Osman, M. B., Das, S. B., Trusel, L. D., Evans, M. J., Fischer, H., Grieman, M. M., Kipfstuhl, S., McConnell, J. R., and Saltzman, E. S.: Industrial-era decline in subarctic Atlantic productivity, Nature, 569, 551, 2019.

Rafter, P., Bagnell, A., Marconi, D., and Devries, T.: Global trends in marine nitrate N isotopes from observations and a neural network-based climatology, Biogeosciences, 16, 2617–2633, https://doi.org/10.5194/bg-16-2617-2019, 2019.

Roshan, S. and DeVries, T.: Efficient dissolved organic carbon production and export in the oligotrophic ocean, Nature communications, 8, 2036, 2017.

Royer, S.-J., Galí, M., Mahajan, A., Ross, O. N., Pérez, G., Saltzman, E. S., and Simó, R.: A high-resolution time-depth view of dimethylsulphide cycling in the surface sea, Scientific reports, 6, 32 325, 2016.

Saltzman, E., King, D., Holmen, K., and Leck, C.: Experimental determination of the diffusion coefficient of dimethylsulfide in water, J GEOPHYS RES-OCEANS, 98, 16 481–16 486, 1993.

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

Schmidtko, S., Johnson, G. C., and Lyman, J. M.: MIMOC: A global monthly isopycnal upper-ocean climatology with mixed layers, J. Geophys. Res.: Oceans, 118, 1658–1672, 2013.

Simó, R.: Production of atmospheric sulfur by oceanic plankton: Biogeochemical, ecological and evolutionary links, Trends in Ecology and Evolution, 16, 287–294, https://doi.org/10.1016/S0169-5347(01)02152-8, 2001.

Simó, R. and Dachs, J.: Global ocean emission of dimethylsulfide predicted from biogeophysical data, Global Biogeochem. Cycles, 16, 26–1, 2002.

Simó, R. and Pedrós-Alió, C.: Short-term variability in the open ocean cycle of dimethylsulfide, Global Biogeochem. Cycles, 13, 1173–1181, 1999.

Stefels, J.: Physiological aspects of the production and conversion of DMSP in marine algae and higher plants, J. Sea Res., 43, 183–197, https://doi.org/10.1016/S1385-1101(00)00030-7, 2000.

Stefels, J., Steinke, M., Turner, S., Malin, G., and Belviso, S.: Environmental constraints on the production and removal of the climatically active gas dimethylsulphide (DMS) and implications for ecosystem modelling, Biogeochemistry, 83, 245–275, 2007.

Stefels, J., Carnat, G., Dacey, J. W., Goossens, T., Elzenga, J. T. M., and Tison, J.-L.: The analysis of dimethylsulfide and dimethylsulfoniopropionate in sea ice: Dry-crushing and melting using stable isotope additions, Marine chemistry, 128, 34–43, 2012.

Sunda, W. G., Kieber, D., and Kiene, R. P.: An antioxidant function of DMSP and DMS in marine algae, Nature, 418, 317–320, https://doi.org/10.1038/nature00851, 2002.

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

Thomas, D. and Dieckmann, G.: Antarctic sea ice–a habitat for extremophiles, Science, 295, 641–644, 2002.

Toole, D., Slezak, D., Kiene, R., Kieber, D., and Siegel, D.: Effects of solar radiation on dimethylsulfide cycling in the western Atlantic Ocean, DEEP-SEA RES PT I, 53, 136–153, 2006.

Toole, D. A. and Siegel, D. A.: Light-driven cycling of dimethylsulfide (DMS) in the Sargasso Sea: Closing the loop, Geophys Res Lett, 31, 2004.

Vallina, S. M. and Simó, R.: Strong Relationship Between DMS and the Solar Radiation Dose over, Science, 315, 506–509, https://doi.org/10.1126/science.1133680, 2007.

Vogt, M., Vallina, S. M., Buitenhuis, E. T., Bopp, L., and Le Quéré, C.: Simulating dimethylsulphide seasonality with the Dynamic Green Ocean Model PlankTOM5, J. Geophys. Res., 115, C06 021, https://doi.org/10.1029/2009JC005529, 2010.

Wang, S. and Moore, J. K.: Incorporating Phaeocystis into a Southern Ocean ecosystem model, J. Geophys. Res., 116, C01 019, https://doi.org/10.1029/2009JC005817, 2011.

Wang, S., Elliott, S., Maltrud, M., and Cameron-Smith, P.: Influence of explicit Phaeocystis parameterizations on the global distribution of marine dimethyl sulfide, J. Geophys. Res.: Biogeosciences, 120, 2158–2177, https://doi.org/10.1002/2015JG003017, 2015.

Wang, W.-L., Moore, J. K., Martiny, A. C., and Primeau, F. W.: Convergent estimates of marine nitrogen fixation, Nature, 566, 205–211, 2019.

Webb, A. v., van Leeuwe, M., den Os, D., Meredith, M., Venables, H., and Stefels, J.: Extreme spikes in DMS flux double estimates of biogenic sulfur export from the Antarctic coastal zone to the atmosphere, Scientific reports, 9, 1–11, 2019.

Yang, G. P., Liu, X. T., Li, L., and Zhang, Z. B.: Biogeochemistry of dimethylsulfide in the South China Sea, Journal of Marine Research, 57, 189–211, https://doi.org/10.1357/002224099765038616, 1999.

Table 1. Results of linear regression models. The R^2 values are for log transformed, and normalized data as described in the text.

	in situ data			PMEL			PMEL+NAAMES		
Parameter	R^2	Slope	No.	R^2	Slope	No.	R^2	Slope	No.
DMSPt	0.41	0.77	4,061	-	-	-	-	-	-
Chl a[1]	0.21	0.43	10,404	0.09	0.30	81,767	0.09	0.29	88,516
MLD	-	-	-	0.06	-0.25	81,646	0.07	-0.26	88,214
PAR	-	-	-	0.07	0.26	82,137	0.09	0.29	88,923
SST	$\sim$0	-0.01	69,196	0.02	-0.12	82,770	0.01	-0.12	89,556
SSS	$\sim$0	-0.08	69,196	0.01	-0.10	82,759	0.02	-0.13	89,545
DIP	-	-	-	0.01	0.11	81,868	0.02	0.12	88,654
DIN	-	-	-	0.01	0.10	79,083	$\sim$0	0.09	85,865
SiO	-	-	-	0.04	0.19	81,813	0.04	0.20	88,599

Table 2. Annually-averaged zonal mean DMS flux (Tg S/yr) for this study (W20), Lana et al. (2011) (L11), Simó and Dachs (2002)(SD02), Vallina and Simó (2007)[VS07], and Galí et al. (2018)[Gali18] for their four parameterization models. L11, SD02, VS07, and Gali18 are computed with the Nightingale et al. (2000) parameterization of the piston velocity[N00]. Flux in this study is calculated using both the Nightingale et al. (2000)[N00], and Goddijn-Murphy et al. (2012)[GM12], parameterizations. Uncertainties are estimated based on top 10 models with different parameterizations. Errorbars correspond to $\pm 1\sigma$.

Latitude	L11[N00]	SD02[N00]	VS07[N00]	Gali18[N00]	W20[N00]	W20[GM12]
90°-80°N	0.00	0.00	0.00	0.00± 0.00	0.00±0.00	0.00±0.00
80°-70°N	0.08	0.04	0.02	0.02± 0.00	0.05±0.01	0.04±0.01
70°-60°N	0.19	0.11	0.06	0.09± 0.01	0.13±0.01	0.11±0.01
60°-50°N	0.78	0.52	0.30	0.38± 0.04	0.45±0.03	0.35±0.03
50°-40°N	1.16	1.01	0.81	0.73± 0.08	0.79±0.06	0.60±0.05
40°-30°N	1.39	1.64	1.85	1.18± 0.07	1.13±0.05	0.90±0.04
30°-20°N	1.43	1.89	2.84	1.33± 0.02	1.29±0.05	1.07±0.04
20°-10°N	2.60	2.79	4.29	1.96± 0.07	2.12±0.09	1.68±0.07
10°-0°N	2.91	2.64	3.55	1.66± 0.03	2.11±0.10	1.79±0.08
00°-10°S	2.90	2.40	3.54	1.84± 0.01	2.23±0.13	1.91±0.11
10°-20°S	3.42	2.64	4.35	2.05± 0.02	2.41±0.13	1.93±0.11
20°-30°S	2.91	2.26	3.74	1.87± 0.02	1.93±0.12	1.56±0.10
30°-40°S	2.91	2.42	3.00	2.19± 0.08	2.20±0.19	1.71±0.14
40°-50°S	2.70	2.19	2.18	2.07± 0.14	2.19±0.16	1.51±0.11
50°-60°S	1.67	1.00	0.10	0.76± 0.07	1.01±0.07	0.67±0.05
60°-70°S	0.18	0.08	0.08	0.04± 0.00	0.09±0.01	0.06±0.01
70°-80°S	0.00	0.00	0.00	0.00± 0.00	0.00±0.00	0.00±0.00
80°-90°S	0.00	0.00	0.00	0.00± 0.00	0.00±0.00	0.00±0.00
Total	27.23	23.64	31.59	18.18± 0.52	20.12±0.43	15.89±0.34

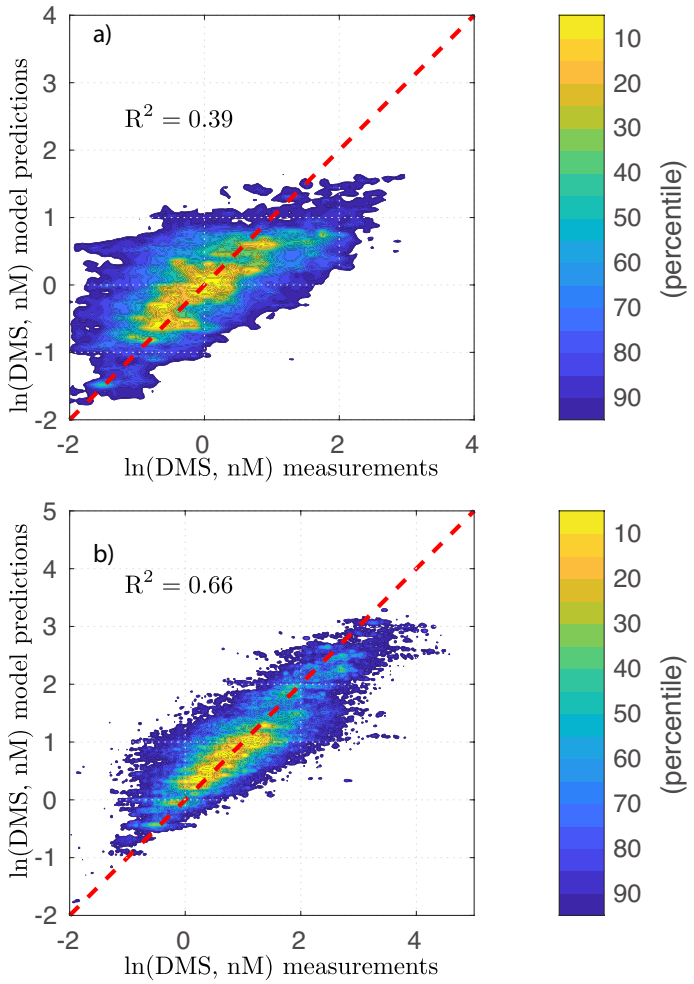

Figure 1. Model versus observation plots on logarithmic scale: (a) multilinear regression model; (b) artificial neural network model. The color indicates the fraction of the joint distribution explained as a percentile that falls within a region of concentration space.

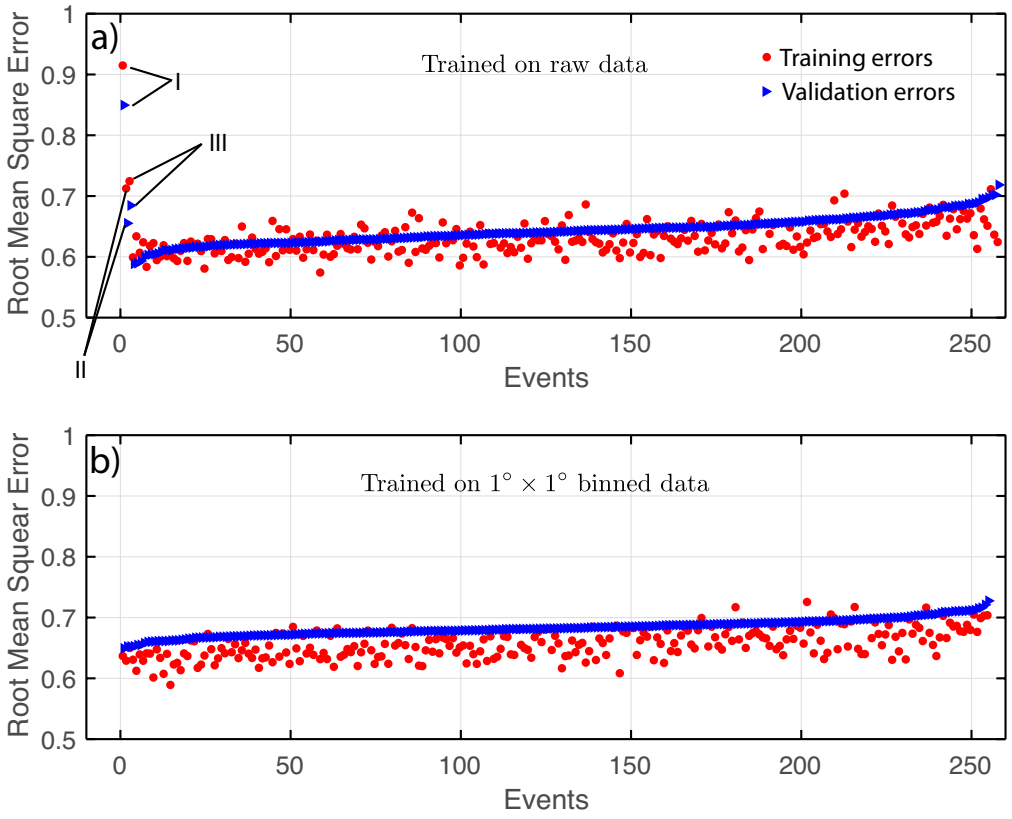

Figure 2. Parameter sensitivity tests on raw and binned data. (a) Root mean square error on logarithmic scale for the model trained using raw data; (b) Root mean square error on logarithmic scale for the model trained using binned data . The time and location parameters are tested separately without combining with environmental parameters as shown in the upper panel, (I) with only location parameters; (II) with location and day of year parameters; and (III) with location, day of year, and time of day parameters. The model with three location parameters (I) has a root mean square error on natural logarithmic scale of ∼0.83, which decreases to ∼0.65 by adding sampling day of year parameters (II), however, increases to ∼0.67 by adding sampling time parameters (III). We, therefore, do not include sampling time parameters in the following tests. We tested every combination of the eight parameters (PAR, MLD, SST, SAL, Chl *a*, DIP, DIN, and SiO), which in total are 255 tests.

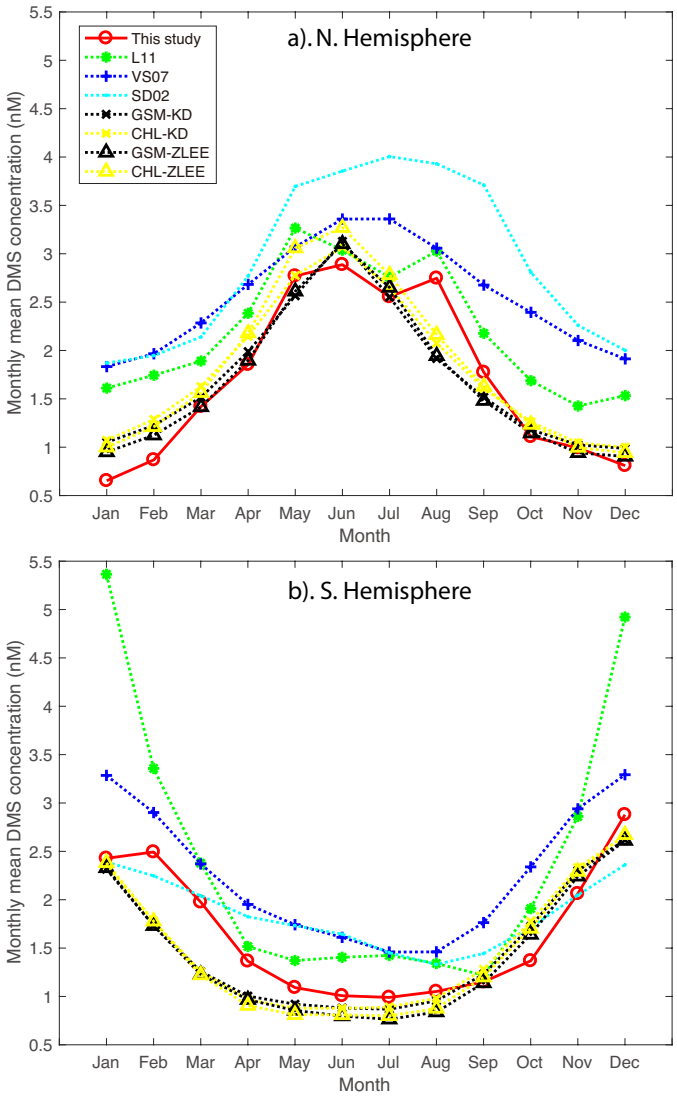

Figure 3. Comparisons of monthly mean DMS concentrations to previous studies (Simó and Dachs, 2002; Vallina and Simó, 2007; Lana et al., 2011; Galí et al., 2018). L11, SD02, and VS07 are self-explanatory. GSM-KD, CHL-KD, GSM-ZLEE, and CHL-ZLEE are the four model results from Galí et al. (2018).

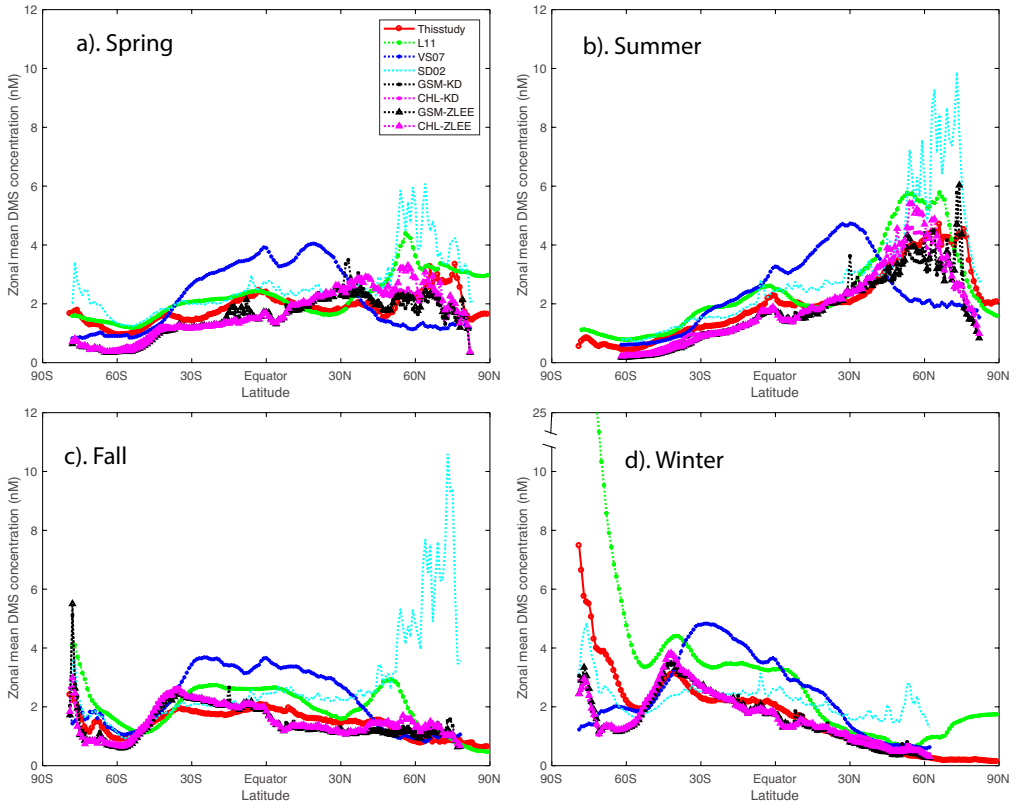

Figure 4. Comparisons of zonally mean DMS concentrations to previous studies (Simó and Dachs, 2002; Vallina and Simó, 2007; Lana et al., 2011; Galí et al., 2018). L11, SD02, and VS07 are self-explanatory. GSM-KD, CHL-KD, GSM-ZLEE, and CHL-ZLEE are the four model results from Galí et al. (2018).

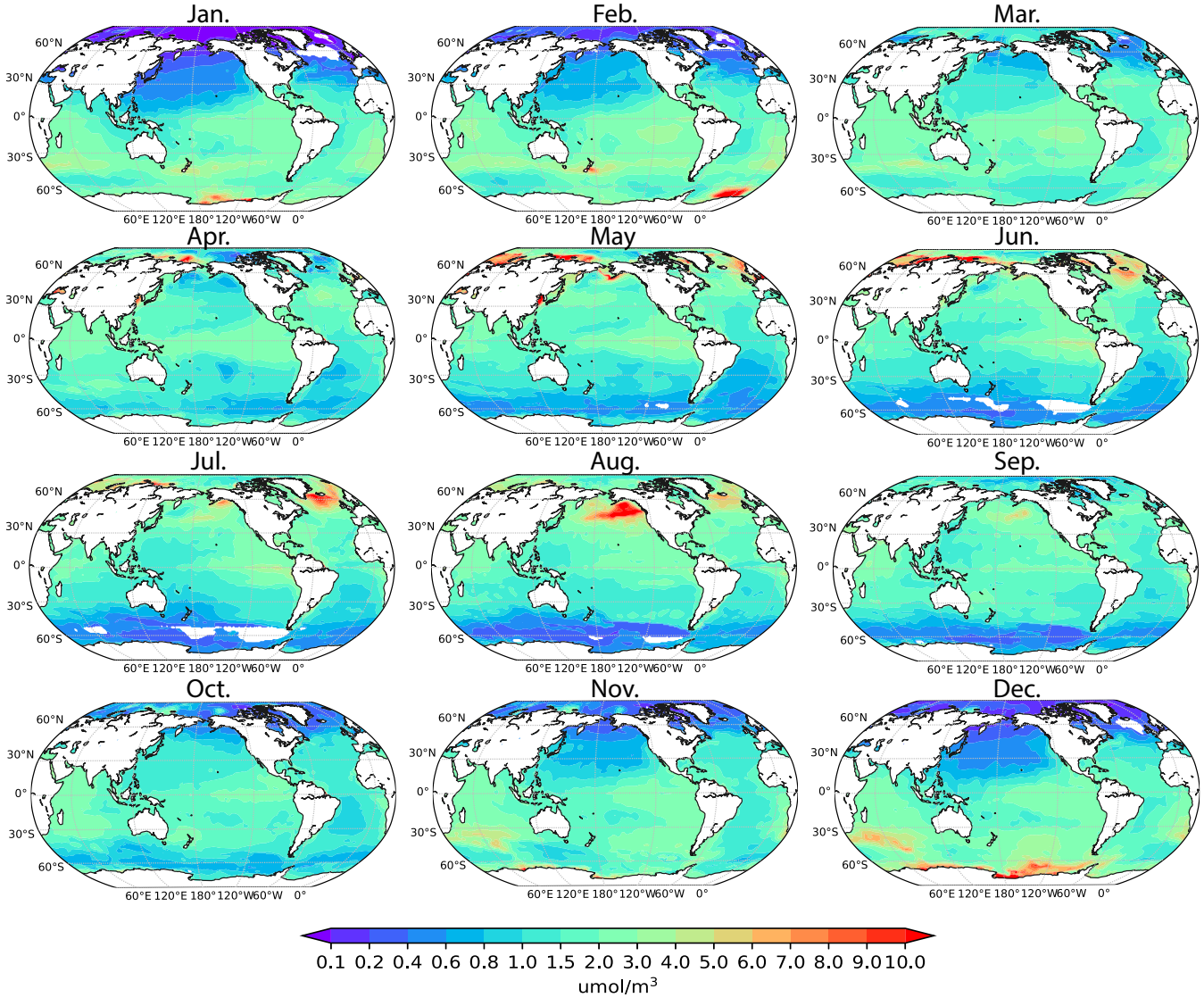

Figure 5. Monthly DMS concentration (μmol m^{-3}) estimated based on artificial neural network.

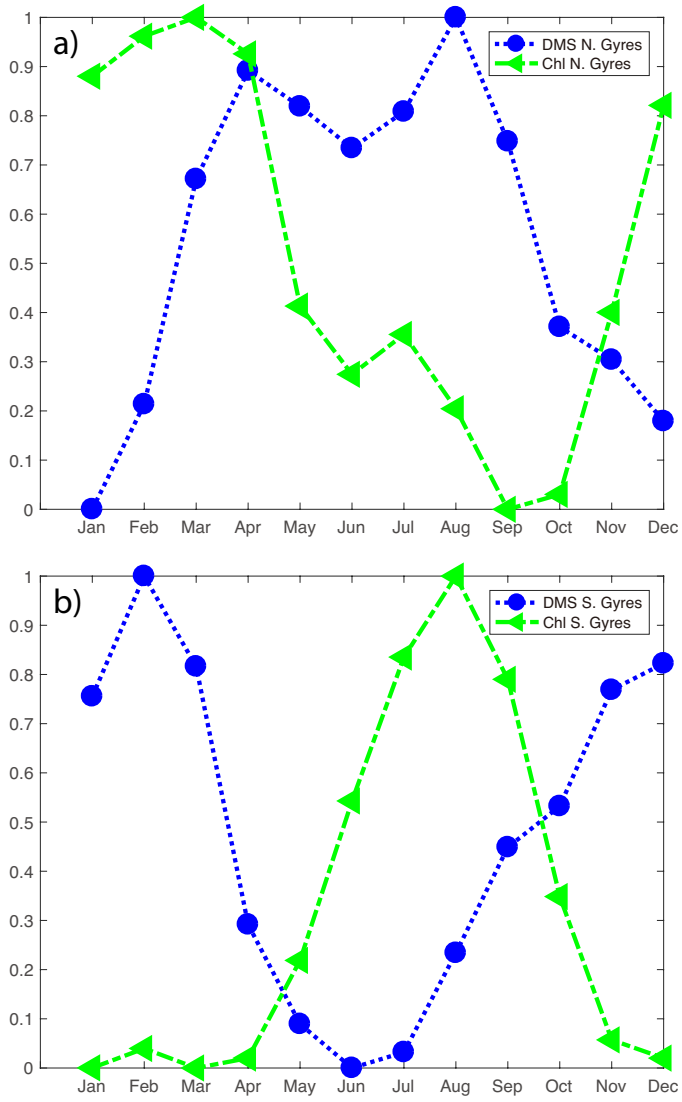

Figure 6. Distributions of monthly mean DMS and Chl-*a* concentrations for N. and S. hemisphere gyres. The gyres are defined as regions between 30° and equator where annually mean DIP concentration is below 0.2 μM. Monthly mean concentrations are normalized to the range of 0 to 1.

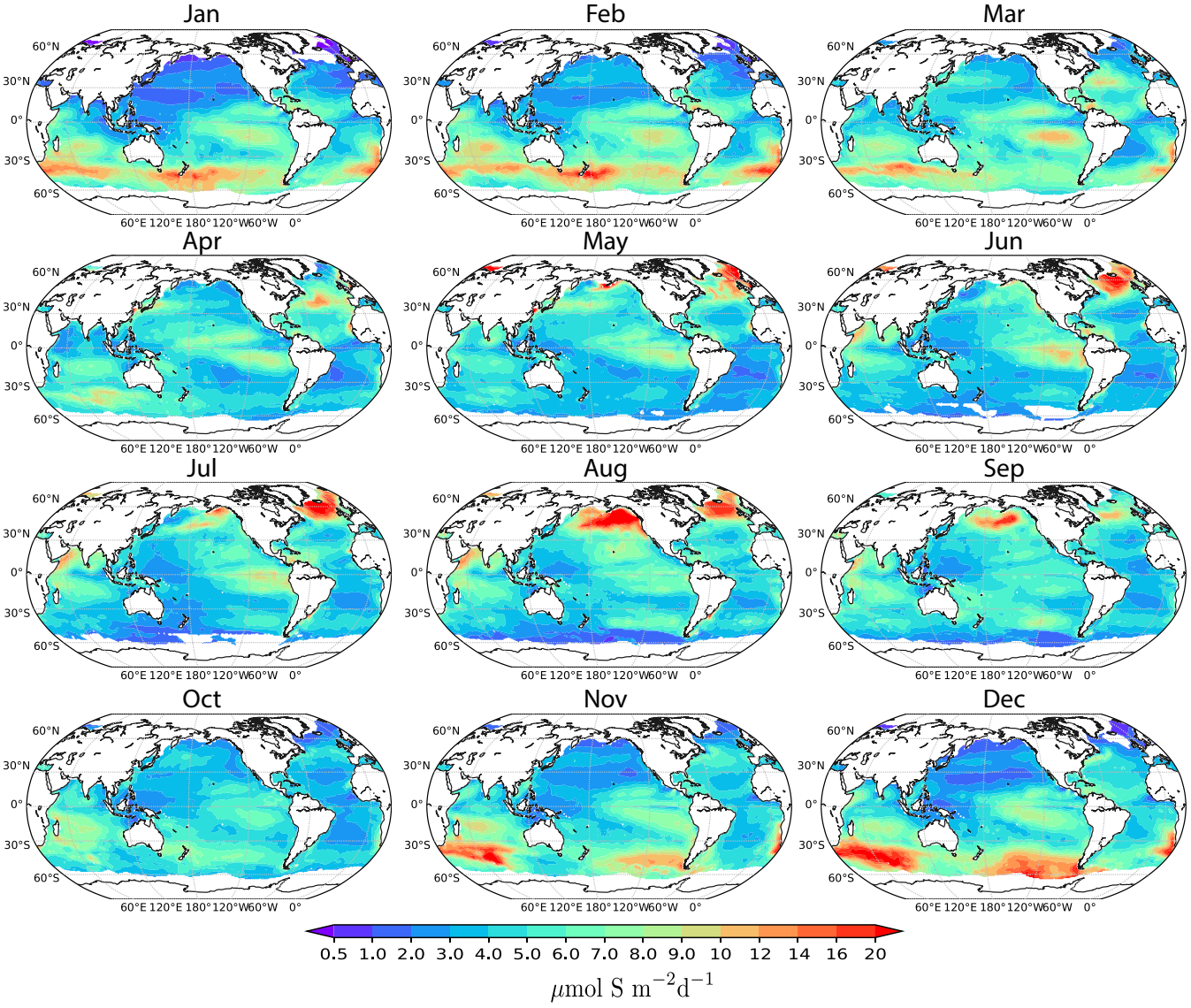

Figure 7. Monthly DMS flux (μmol S m^{-2} day^{-1}) calculated based on DMS climatology estimated from the ANN model and Goddijn-Murphy et al. (2012) flux parameterization.

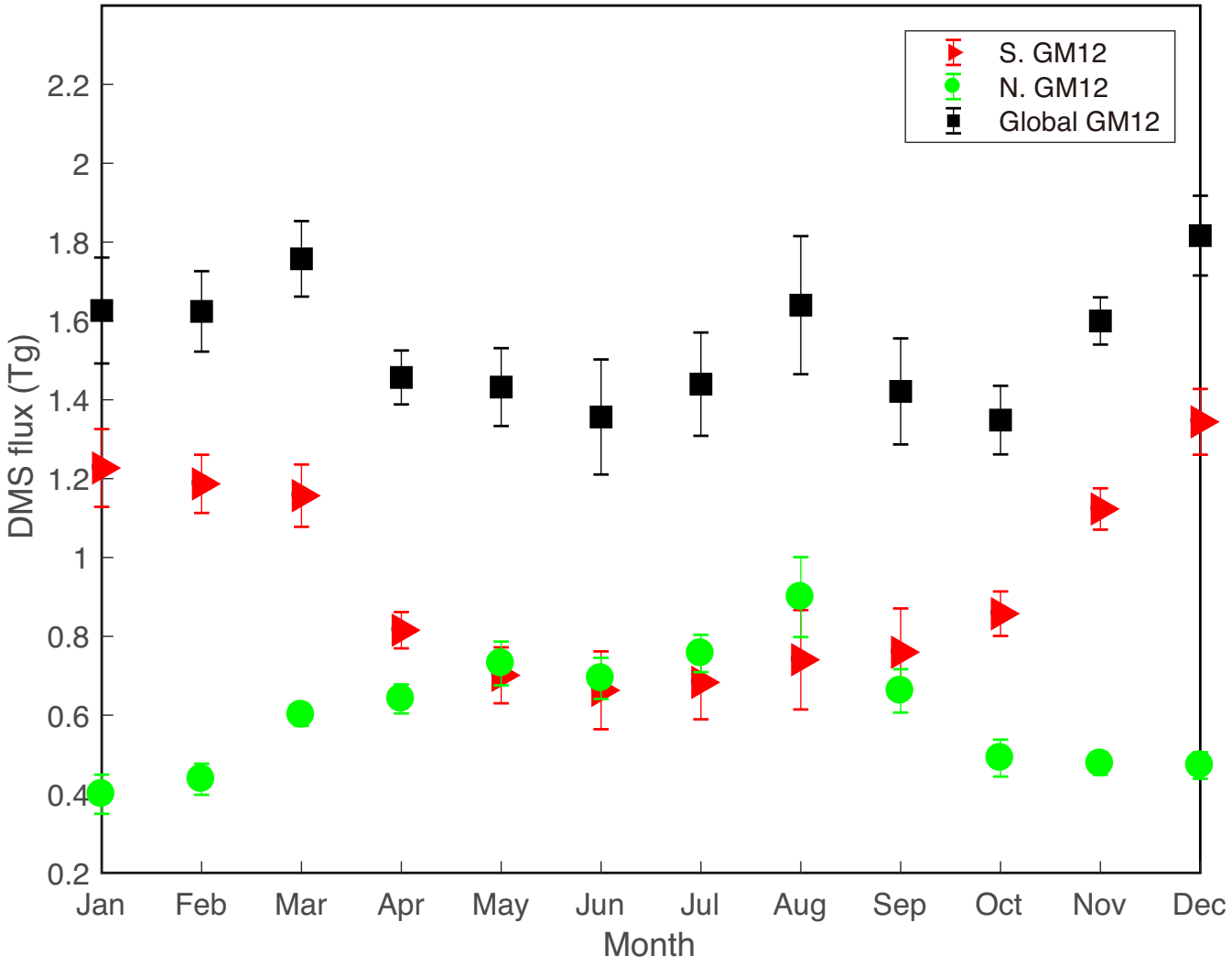

Figure 8. Area and month integrated DMS sea-to-air flux (Tg S month^{-1}) based on GM12 parameterization. Red triangles represent monthly mean flux of the Southern hemisphere, green dots represent monthly mean flux of the Northern hemisphere, and black squares represent globally monthly mean flux. Uncertainties are estimated based on top 10 models with different parameter combinations. Errorbars correspond to $\pm 1\sigma$.

Table A1. DMS and ancillary data sources.

Variables	Sources	units	References
DMS[1]	http://saga.pmel.noaa.gov/dms/	nM	(Kettle et al., 1999)
DMS[2]	NAAMES	nM	(Behrenfeld et al., 2019)
Chl	https://oceandata.sci.gsfc.nasa.gov/SeaWiFS/	μg L^{-1}	(NASA, 2018)
MLD	https://www.pmel.noaa.gov/mimoc/	m	(Schmidtko et al., 2013)
PAR	https://oceancolor.gsfc.nasa.gov/atbd/par/	Einsteins m^{-2} d^{-1}	(Frouin et al., 2012)
WSP	https://podaac.jpl.nasa.gov/dataset	m s^{-1}	(NASA, 2012)
SST	WOA2013	C	(Garcia et al., 2013)
SSS	WOA2013	psu	(Garcia et al., 2013)
DIP	WOA2013	μM	(Garcia et al., 2013)
DIN	WOA2013	μM	(Garcia et al., 2013)
SiO	WOA2013	μM	(Garcia et al., 2013)
ICE	CESM model	-	(Wang et al., 2019)

[1] Data from the online database. [2] New data from the North Atlantic Aerosol and Marine Ecosystems experiment.

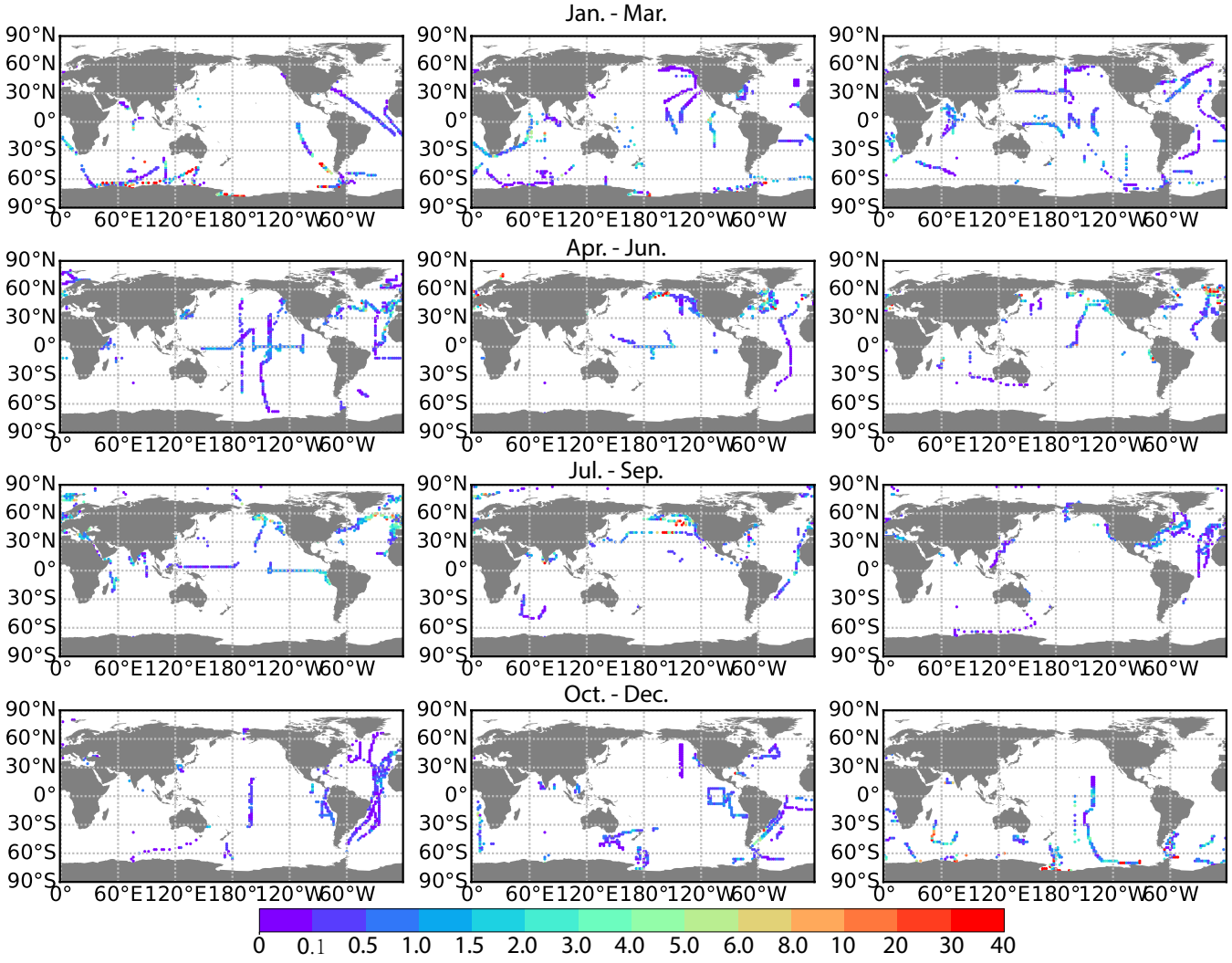

Figure A1. Distribution of DMS observations partitioned into each month. The color indicates DMS concentration (nM).

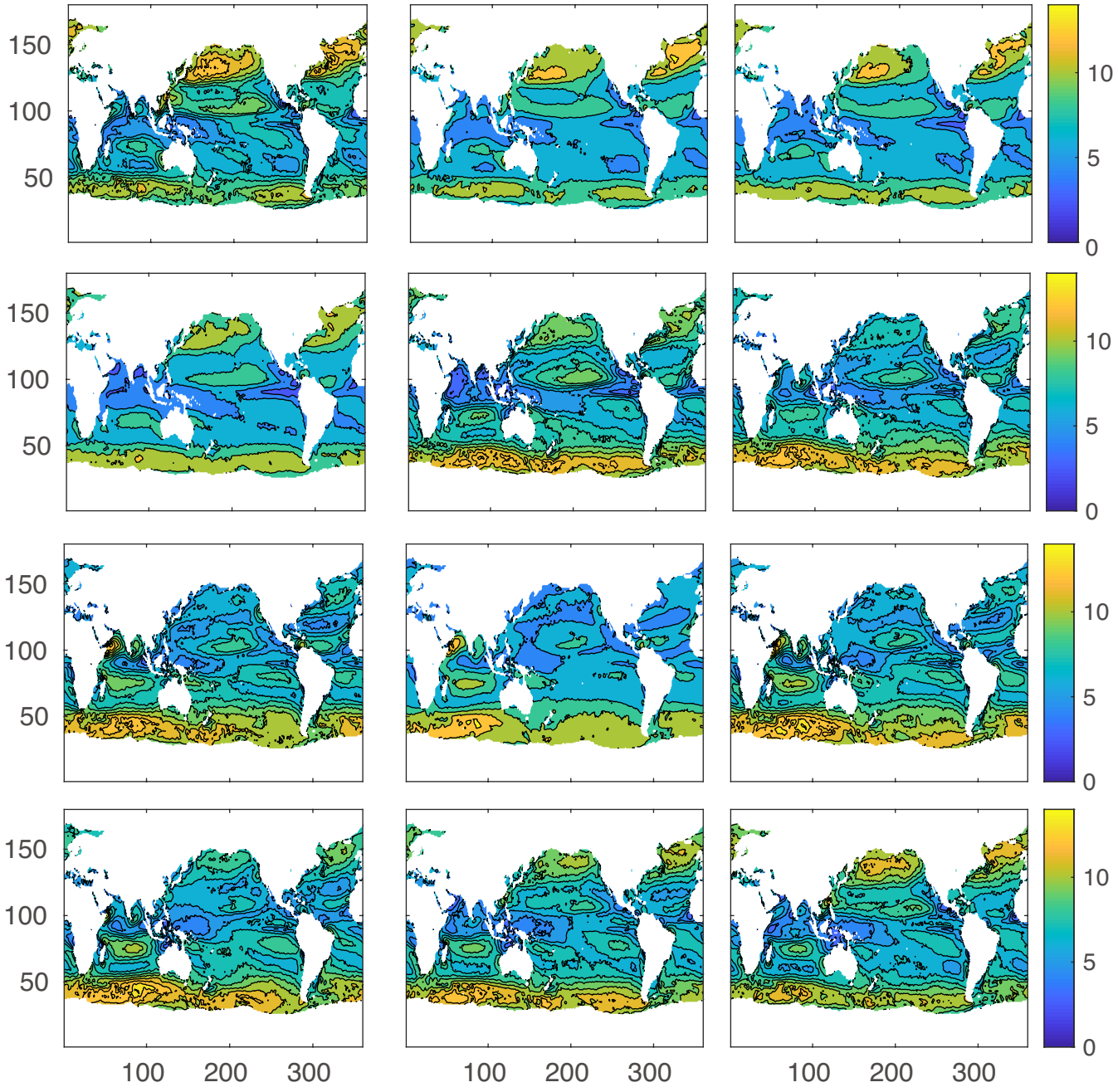

Figure A2. Climatological wind speed (m s^{-1}).