Manuscript "Evaluation of ocean dimethylsulfide concentration and emission in CMIP6 models" by Bock and coauthors

Review by Martí Galí

GENERAL COMMENTS

The article by Bock an coauthors gives a comprehensive overview of sea-surface DMS concentrations and sea-air fluxes in four CMIP6 models, which feature DMS parametrizations of different complexity (two diagnostic and the other two prognostic). The models also differ in their coupling (or lack thereof) between marine DMS emission and atmospheric chemistry and aerosol/cloud radiative forcing, although this aspect is not dealt with in the discussion. The main conclusions of the article are (1) the uncertainty in present-day model estimates, with spatial patterns and seasonal cycles that differ from observational products in many oceanic regions, and (2) the diverging trends in global DMS emissions in end-of-century projections. These divergences reflect the different factors that drive DMS concentrations in each model, rather than the complexity of the parametrizations. The findings reported by Bock and coworkers will be useful to advance the modelling of marine DMS cycling because they highlight the mixed success in representing DMS cycling appropriately in CMIP numerical models and, more generally, the gap between experimental/observational knowledge and model parametrizations.

Below I summarize my main criticisms, with the hope that they will help improve the manuscript:

1. On the potential overestimation of global DMS concentration by the Lana et al. (2011) climatology, and the failure to capture extremes. The comparison between ESM results, global DMS climatologies based on statistical relationships and EO datasets (G18, W20), and the L11 climatology based on objective interpolation of in situ data, indicates greater consistency between the former two. The authors conclude that L11 overestimates DMS globally, a conclusion that is supported by previous works (Tesdal et al. 2016; Galí et al, 2018). Although the conclusions of the latter studies still hold over many regions, here I would like to note that all model and statistical representations of sea-surface DMS fields fail to account for extreme DMS concentrations. As the authors note, extreme concentrations were removed to compute the L11 climatology, although they were not systematically identified as measurement artefacts. Webb et al. (2019, SciRep) showed the importance of extreme DMS events in Antarctica. Recently, Bell et al. (2020, Frontiers) showed that both the L11 climatology and the observations-based G18 and W20 diagnostics failed to capture high-DMS events in the North Atlantic. In the light if these findings, the message that L11 overestimates DMS globally should be nuanced. Obviously, there is a dichotomy between the ability to capture the mean state and the extremes, and the latter might play an important role in ocean-atmosphere interactions.

2. Global NPP-DMS relationship. Understanding the global NPP-DMS relationship would be useful to place emergent constraints on present-day and future DMS emission. The attempt made in this article to pinpoint this relationship is very welcome, but the discussion of the underlying factors is poor. First, I would not expect that studies covering small spatiotemporal scales (e.g. those cited in L601) could give relevant insights into the NPP-DMS relationship over multiyear periods at the biome scale. Perhaps the work that addressed this issue more explicitly was that of Kameyama et al. (2013, GRL. Strong relationship between dimethyl sulfide and net community production in the western subarctic Pacific); but they related DMS to NCP, not NPP. I also recommend the work of Osman et al. (2019, Nature. Industrial-era decline in subarctic Atlantic productivity). Second, the discussion of the NPP-DMS relationship disregards the complexity of food-web and abiotic processes that control DMS concentrations. The framework proposed by Galí and Simó (2015, GBC) could be useful to understand the contribution of different factors to DMS variability in prognostic models.

3. Relevant literature. I suggest that authors to dig deeper into the non-modelling literature, which is connected to the point above. The article currently gives the impression that the authors are not familiar enough with some aspects of the DMS and DMSP biogeochemistry because imprecise informations are scattered in the text (see specifics). Suggestions of relevant articles can be found through this review. In addition, I strongly suggest the authors to pay more attention to the body of modelling literature produced by the group of Elliott and colleagues:
Wang et al. (2015, JGR). Influence of explicit *Phaeocystis* parameterizations on the global distribution of marine dimethyl sulfide.
Wang et al. (2018, Biogeochemistry). Influence of dimethyl sulfide on the carbon cycle and biological production.
Xu et al. (2016, JGR). DMS role in ENSO cycle in the tropics.
Another important article that explains the different ability of models to decouple DMS from phytoplankton biomass is:
Le Clainche et al. (2010, GBC). A first appraisal of prognostic ocean DMS models and prospects for their use in climate models

4. Extrapolation of DMS emission to an ice-free Arctic summer. A correction is needed here. In the G19 paper we warned that extrapolations towards a 100% ice-free Arctic summer shouldn't be made using the pan-Arctic linear regression between DMS emission and ice-free extent. This would amount to assuming that the % contribution from each subregion will not change in the future, which we know is not true. The Atlantic sector can't contribute much more than it presently does because it already is mostly ice free; on the contrary, the Central Arctic is far from being ice-free in summer and has the lowest DMS concentrations (Uhlig et al. 2019, Frontiers), thus the lowest potential for an increase in emissions. Of course, this is illustrated by the fact that the sums of regional extrapolations do not equal the pan-Arctic extrapolation in Table 9. I argue that, to compute the range of future ice-free Arctic DMS emissions, only the sum of regional contributions in each model should be used. I also recommend the study of Hayashida et al. (2020, GBC. Spatiotemporal Variability in Modeled Bottom Ice and Sea Surface Dimethylsulfide Concentrations and Fluxes in the Arctic During 1979–2015) who used a regional model with higher resolution to estimate contemporary DMS emission from the Arctic.

SPECIFIC COMMENTS

L23: perhaps a "by-product of microbial food webs" is more accurate, given the important roles of heterotrophic processes (grazing, bacterial catabolism) in DMS production. In other words, DMS production can't be understood from phytoplankton processes alone (which is different from saying that it cannot be predicted from phytoplankton variables…).

L25: please don't forget MSA, a product of the addition pathway that can be produced in relevant proportions compared to sulfuric acid.

L30: the original CLAW hypothesis paper (Charlson et al., 1987, Nature) should be given credit here. Regarding the relationship between DMS and downwelling irradiance, please consider citing Vallina and Simó (2007, Science).

L87: This is inaccurate. I suggest: "Dissolved DMSP is then converted to DMS with yields that increase with bacterial nutrient stress". Note: PISCES assumes all DMS production arises from dissolved DMSP, unlike other models, and contrary to observations. To compensate for this, PISCES requires "bacterial" DMS yields that range between 40 and 60% (Belviso et al., 2012), which are clearly too high (Galí and Simó, 2015).

Table 1: In my opinion, this table should cite the original articles where the prognostic or diagnostic DMS models were described. In the case of the diagnostic ones it is very simple:
- MIROC uses Aranami and Tsunogai (2004)
- UKESM uses Anderson et al. (2001)
This may be more difficult in the case of prognostic models that have seen incremental development. In this case, I suggest citing in this table the most recent papers where each prognostic model was described.

Section 2.1.1: I did a little research on the prognostic sulfur modules.
- PISCES: As the authors describe in this section, the original sulfur module of Bopp et al. (2008) was updated by Belviso et al. (2012) based on the model PlankTOM5 of Vogt et al. (2010). This detail should be included.

L187: This is true but the explanation is inaccurate. What is right-skewed is the distribution of satellite-retrieved Chl concentration matched to the DMS database, compared to the global distribution (PDF) of satellite Chl. As the authors correctly point out in the following sentence, this is related to preferential sampling of productive waters that probably had higher-than-average DMS. Ultimately, this could partially explain the right skewness of the in situ measurements when compared to the global DMS fields estimated with the G18 algorithm. But the latter may also suffer from biases caused by the fitted equations and satellite observations.

L217: besides sea ice, what limits satellite ocean colour measurements (passive radiometry) at high latitudes in winter is the low solar elevation. In December, no reliable observations are available north of about 48 degrees, in November and January the boundary is slightly above 50 degrees, etc.

L306-308: I disagree here. For unknown reasons, models struggle to capture high DMS in regions like the NE Pacific (station PAPA) and around Antarctica. In the NE Pacific, occurrence of very high DMS (often >15 nM) in late summer has been extensively documented by the Line P program and by many studies from Philippe Tortell's group. The article of Steiner et al. (2012, Biogeochemistry) explored potential reasons. This was also discussed by Galí et al. 2018. In the Southern Ocean, *Phaeocystis antarctica* blooms can results in seawater DMS of several tens nM, e.g. del Valle et al. (2009, L&O), Webb et al. (2019, SciRep). Highest DMS concentrations, many of which measured around Antarctica, were removed by Lana et al. (2011) before computing their climatology. So there's a general failure at capturing extreme DMS concentrations under certain conditions. See general comments.

L324: Note that the high DMS in winter and spring at high northern latitudes in L11 could be an interpolation artifact, as explained by Hayashida et al. (2020, GBC). This is relevant to BPLR panel in Fig. 5.

L349: "six biomes", but only four are listed. Perhaps specify that N and S polar and westerlies biomes are treated separately (but not trades or coastal ones).

L362: Well, low correlations are expected in areas where DMS has low seasonal amplitude. This also applies to the paragraphs below. OK, this was answered in L383.

L372: Looking at Fig. 5 province by province, and at Table 5, my first impression is that NORESM usually does better at capturing seasonal cycles, except for the trades biome. Also, my gut feeling is that this model better captures summertime DMS maxima at subtropical to temperate latitudes (the DMS summer paradox; check Simó & Pedrós-Alió 1999, Nature; LeClainche et al. 2010, GBC; Galí and Simó 2015, GBC), at least in regions I know best (4-NADR, 6-NASW, 16-MEDI).

Figure 7: please specify somewhere (figure or caption) that CAMS uses the N00 parameterization.

L480-481: I tend to agree with Tesdal here. See general comments.

L565: “regions or processes”

L585: “modelling works”. Otherwise a large body of non-modelling literature would be disregarded.

L601-606: The discussion of the relationship between NPP and DMS is poor, and misses some relevant literature. DMS production is a food web process, not just a phytoplankton process. See the general comments.

Figure 14: Axis labels are too small, and some horizontal reference lines (or a grid) would be very useful to guide comparisons among models.

L641-648: It is important to note that the non-Atlantic sector includes the Siberian shelves, which seem to be quite productive owing to nutrient inputs from large rivers, recycling on the shelves and coastal erosion (Terhaar et al., 2021, NatComm). However, the satellite data used in G19 may be biased high in the Siberian shelves due to optical interference of continental materials (which was also pointed out by Hayashida et al., 2020). So I would dare to say that uncertainty in satellite DMS is much higher in the non-Atlantic sector, and that ESMs possibly struggle to capture the biogeochemical functioning in shallow Arctic seas, due to both too-low resolution and non-represented processes.

L665: Since emissions arise mostly from ice-free areas in both satellite and ESM asessments, additional factors must be invoked to explain scatter in the ice extent vs. DMS emission relationships. I am pretty sure that lower R2 in models results from too-low interannual variability in models compared to satellite observations. In G19, we pointed out that after 2011 interannual variability was controlled mostly by ocean productivity, not ice extent. This was further analyzed, and confirmed, by Lewis et al. (2020, Science).

Figure 15 and the related analysis: see general comments.

L687: Please beware that Le Clainche et al. (2010) did a DMS model intercomparison.

L694: agree on sampling biases, disagree on global overestimation. See general comments.

L701: ...which has been known for a long time (see Le Clainche et al., 2010 and references thererin)

Typos, technical corrections
L23: space after (DMS)
L26: something is missing: “formed DMS”
L116: open parentheses before “Simó and Dachs”
L212-215: please consider breaking this sentence with a period somewhere.
L277: “maxima”, not “maximums”
L376: hypotheses
L576: “at play”?