

The authors came up with data-driven models to simulate and predict the methanesulfonic acid in Arctic aerosol based on the observed data from 4 sites, and compared the models with traditional numerical models. I recommend publication after the following issues are addressed.

We thank the reviewer for their comments and suggestions. We have addressed each comment below with review comments in black, author response in blue, and additions to the original text in red. We have indented the author's response for clarity. Line numbers given in the author's response refer to lines in the revised manuscript.

1. In "Abstract" Sec., line 24 and line 30, input features consider ambient conditions/ sources, chemical processing, and MSA removal. These two kinds of features do not include each other, it's better to make the expression unified.

To clarify, the input features for our data-driven models consider the ambient conditions experienced during air mass transport, rather than the in situ conditions measured at the measurement location. These include meteorological, oceanic, and geographical variables, listed in Table 2. In our interpretation of the output from these data-driven models, we identified three groups of features selected by the models to be important (sources, chemical processing, and removal). We have adjusted the text to clarify this.

Lines 24-25: In our approach, we create input features that consider the ambient conditions **experienced** during atmospheric transport (e.g., **DMS emission**, temperature, radiation, cloud cover, **precipitation** etc.)

Line 31-32: The data-driven models selected features **which can be grouped into three categories**, the sources, chemical processing, and removal of MSA_p.

2. In "Introduction" Sec., lines 58-60, "which is enzymatically cleaved to produce DMS... in the atmosphere", better to write as "which is enzymatically cleaved to produce DMS, acting as the main source of marine atmosphere/aerosol".

We have amended the text according to the reviewer's suggestion.

Lines 61-62: Arctic marine phytoplankton and algae produce dimethylsulfoniopropionate as an osmoprotectant (Yoch, 2002), which is enzymatically cleaved to produce **seawater** DMS (Andreae, 1990; Kettle et al., 1999), **which** is the main source of marine biogenic sulfur in the atmosphere (Hulswar et al., 2022; Lana et al., 2011).

3. In "Introduction" Sec., lines 68-69, the references should decrease to 2-3 classic or newest papers.

There is still uncertainty regarding the exact formation mechanisms of DMS oxidation into MSA, therefore, we list several modelling studies (Chen et al., 2018; Fung et al., 2022; von Glasow and Crutzen, 2004; Hoffmann et al., 2016) utilizing chemical transport and box models, a study from the Arctic utilizing observations and box modelling (Kecorius et al., 2023), an observational study from the Southern Ocean (Baccarini et al., 2021), and a study combining chamber measurements with box modelling (Wollesen de Jonge et al.,

2021). Therefore, we feel these references give the reader ample resources regarding DMS oxidation into MSA from several different perspectives.

4. In “Introduction” Sec., line 74, a repetition of lines 63-64, also line 277, please make it more concise.

The information on line 63-64 of the original manuscript describes the lifetime of gaseous DMS while the information on line 74 describes the lifetime of particulate MSA. These are two important pieces of information which was used in our decision of the duration of the air mass history on line 308-309 of the revised manuscript. We feel this information is vital to the reader and motivation for our methodological decision making, therefore we will keep this information as is.

5. In “Introduction” Sec., paragraph 2, the description of the relationship between DMA and MSA in the Arctic marine environment is presented in about 30 rows. Authors should adjust the proportion of this in the introduction, meanwhile, pay more attention to the description of numerical models and data-driven models applied in marine research and comparison of them, especially in the Arctic.

We have shortened the paragraph on DMS and MSA although Reviewer 2 wanted more detailed information about DMS oxidation processes to MSA, which we have added. We added more information about numerical modelling of DMS and MSA in the Arctic in the Introduction.

Lines 125-138: Many of these shortcomings are due to lack of knowledge concerning natural processes including rates and spatial distribution of DMS emission, oxidation mechanisms, and cloud processes. Ghahreman et al. (2017) showed that GEOS-Chem overestimated (underestimated) gaseous DMS in summer (spring) in the Canadian archipelago. The overestimation could be attributable to missing aqueous-phase oxidation mechanisms in GEOS-Chem while the underestimation in spring could be due to errors in the DMS source strength (Lana et al., 2011), with missing emissions from melt ponds and marginal ice zones (Gourdal et al., 2018; Hayashida et al., 2017; Mungall et al., 2016). Ghahreman et al. (2021) used the Global Environmental Multi-scale model–Modeling Air quality and Chemistry (GEM-MACH) model to demonstrate that the inclusion of DMS greatly improved the simulated size distribution compared to observations in the Arctic. However, errors in the parameterized nucleation mechanisms led to discrepancies for particles smaller than 50 nm, having implications for cloud formation as aerosols of these sizes have been shown to activate in the Arctic (Leaith et al., 2016). Hoffmann et al. (2021) was able to improve simulations of gaseous MSA in the ECHAM-HAMMOZ model by implementing aqueous-phase oxidation mechanisms on deliquesced particles and by considering the reactive uptake of methanesulfinic acid (MSIA). However, in-cloud processing of MSA is still missing from this model configuration and reactive uptake coefficients are not well parameterized as they depend on aerosol acidity, thus further improvements are required.

6. In “Results” Sec. 3.2, lines 610-622, this paragraph should move to Methods.

We have moved lines 610-617 of the original manuscript to the Sect. 2.7 of the Methods (lines 560-567 of the revised manuscript), we have kept the text on lines 617-622 of the original manuscript in place to state our intentions for this comparison and introduce the section.

7. In “Results” Sec. 3.3, line 743, how to convince readers of the “high accuracy” of max $R^2=0.29$, may be compared to numerical models, it is better to change an appropriate expression of it. Fig.9 is good evidence, maybe move the expression to Fig. 9 discussing.

To clarify, the intention of this text is not to state that our data-driven models have “high accuracy” but that they “struggle to reproduce observed MSA_p with high accuracy”. We can see how this formulation of the text can create confusion. We have amended the text to clarify this.

Lines 807-808: While our data-driven models struggle to accurately reproduce the observed MSA_p (max $R^2 = 0.29$), they can capture the variability (PCC up to 0.77), and they outperform the classic numerical models.

8. In “Results” Sec. 3.3, line 746, there is no performance comparison between numerical and data-driven models in Fig. 4, only data-driven models including RF and AM, please confirm.

We thank the reviewer for this good catch. We have added text to indicate the numerical models are compared against in situ observations in Figs. 3, S2, S3, and S4 and the evaluation metrics for the data-driven models are given in Fig. 4.

Lines 809-811: This is evident from a comparison of the negative R^2 values for the numerical models (Figs. 3, S2, S3, and S4), indicating the numerical models are worse at predicting MSA_p compared to the mean of the observations, versus the evaluation metrics for the data-driven models (Fig. 4).

9. In “Results” Sec. 3.3, line 750, it is confusing to mention Fig. 9 and Fig. 10 here, not to quote the latter figures when discussing the present figure.

We have moved this text to lines 1207-1209 of the revised manuscript to avoid discussing the figures out of order.

Andreae, M. O.: Ocean-atmosphere interactions in the global biogeochemical sulfur cycle, *Mar. Chem.*, 30, 1–29, [https://doi.org/10.1016/0304-4203\(90\)90059-L](https://doi.org/10.1016/0304-4203(90)90059-L), 1990.

Ghahreman, R., Norman, A.-L., Croft, B., Martin, R. V., Pierce, J. R., Burkart, J., Rempillo, O., Bozem, H., Kunkel, D., Thomas, J. L., Aliabadi, A. A., Wentworth, G. R., Levasseur, M., Staebler, R. M., Sharma, S., and Leaitch, W. R.: Boundary layer and free-tropospheric dimethyl sulfide\hack\break in the Arctic spring and summer, *Atmospheric Chem. Phys.*, 17, 8757–8770, <https://doi.org/10.5194/acp-17-8757-2017>, 2017.

Ghahreman, R., Gong, W., Beagley, S. R., Akingunola, A., Makar, P. A., and Leaitch, W. R.: Modeling Aerosol Effects on Liquid Clouds in the Summertime Arctic, *J. Geophys. Res. Atmospheres*, 126, e2021JD034962, <https://doi.org/10.1029/2021JD034962>, 2021.

Gourdal, M., Lizotte, M., Massé, G., Gosselin, M., Poulin, M., Scarratt, M., Charette, J., and Levasseur, M.: Dimethyl sulfide dynamics in first-year sea ice melt ponds in the Canadian Arctic Archipelago, *Biogeosciences*, 15, 3169–3188, <https://doi.org/10.5194/bg-15-3169-2018>, 2018.

Hayashida, H., Steiner, N., Monahan, A., Galindo, V., Lizotte, M., and Levasseur, M.: Implications of sea-ice biogeochemistry for oceanic production and emissions of dimethyl sulfide in the Arctic, *Biogeosciences*, 14, 3129–3155, <https://doi.org/10.5194/bg-14-3129-2017>, 2017.

Hoffmann, E. H., Heinold, B., Kubin, A., Tegen, I., and Herrmann, H.: The Importance of the Representation of DMS Oxidation in Global Chemistry-Climate Simulations, *Geophys. Res. Lett.*, 48, e2021GL094068, <https://doi.org/10.1029/2021GL094068>, 2021.

Hulswar, S., Simó, R., Galí, M., Bell, T. G., Lana, A., Inamdar, S., Halloran, P. R., Manville, G., and Mahajan, A. S.: Third revision of the global surface seawater dimethyl sulfide climatology (DMS-Rev3), *Earth Syst. Sci. Data*, 14, 2963–2987, <https://doi.org/10.5194/essd-14-2963-2022>, 2022.

Kettle, A. J., Andreae, M. O., Amouroux, D., Andreae, T. W., Bates, T. S., Berresheim, H., Bingemer, H., Boniforti, R., Curran, M. A. J., 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, *Glob. Biogeochem. Cycles*, 13, 399–444, <https://doi.org/10.1029/1999GB900004>, 1999.

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 dimethylsulfide concentrations and emission fluxes in the global ocean, *Glob. Biogeochem. Cycles*, 25, GB1004, <https://doi.org/10.1029/2010GB003850>, 2011.

Leaitch, W. R., Korolev, A., Aliabadi, A. A., Burkart, J., Willis, M. D., Abbatt, J. P. D., Bozem, H., Hoor, P., Köllner, F., Schneider, J., Herber, A., Konrad, C., and Brauner, R.: Effects of 20–100 nm particles on liquid clouds in the clean summertime Arctic, *Atmospheric Chem. Phys.*, 16, 11107–11124, <https://doi.org/10.5194/acp-16-11107-2016>, 2016.

Mungall, E. L., Croft, B., Lizotte, M., Thomas, J. L., Murphy, J. G., Levasseur, M., Martin, R. V., Wentzell, J. J. B., Liggio, J., and Abbatt, J. P. D.: Dimethyl sulfide in the summertime Arctic

atmosphere: measurements and source sensitivity simulations, *Atmospheric Chem. Phys.*, 16, 6665–6680, <https://doi.org/10.5194/acp-16-6665-2016>, 2016.

Yoch, D. C.: Dimethylsulfoniopropionate: Its Sources, Role in the Marine Food Web, and Biological Degradation to Dimethylsulfide, *Appl. Environ. Microbiol.*, 68, 5804–5815, <https://doi.org/10.1128/AEM.68.12.5804-5815.2002>, 2002.