Answers to Referee 2

We would like to express our gratitude to the reviewers for their insightful comments on the manuscript. According to the comments, the manuscript has been revised. Please find our replies in the pdf file. In the responses, the reviewer's comments are in black text, and our responses are in blue and *the main text modifications to the revised manuscript are in italics*.

The study of Mansour and colleagues represents a step forward towards the prediction of biogenic sulfur in aerosols, which has climatic and geochemical importance. The authors used several machine learning approaches, each with alternative configurations, to estimate the concentration of the two major atmospheric oxidation products of plankton-made dimethyl sulfide: non-sea-salt sulfate and methanesulfonate. Finally, the best performing model was used to produce daily gridded datasets for these compounds over the North Atlantic Ocean. I found the study methodologically robust and well written, but some issues should be addressed before publication.

General comments

I suggest using nss-SO4, not just SO4, throughout. Abbreviating nss-SO4 may confuse readers because, unlike MSA, SO4 has large anthropogenic and volcanic sources. The same applies to MSA:nss-SO4 ratios.

We agree with the reviewer. The abbreviation has been changed throughout the manuscript (nss-SO$_4^=$ instead of SO$_4$), including the figure axes labels/ captions and table headings/ captions. In the revised manuscript, the following clause has been eliminated:

~~*“Throughout the present study, we abbreviate the nss-SO$_4^{2-}$ concentration as SO$_4$ and MSA concentration as MSA, for simplicity.”*~~

L141: Please provide a quantitative comparison between nss-SO4 and the non-refractory SO4 pool measured with the HR-ToF-AMS, e.g. an indication of the mean absolute and/or relative difference between the two estimates. Just stating they are “approximately equivalent” is not very reassuring. Can the authors exclude the possibility that, in some instances, significant proportions of nss-SO4 are in aerosol fractions not captured by the HR-ToF-AMS?

Our statement is based on the general understanding of the AMS measuring principle. Indeed, the AMS is very sensitive to sulfate in the form of ammonium sulfate, ammonium bisulfate and sulfuric acid (Chen et al., 2019; DeCarlo et al., 2006), which are the main forms under which nss-$SO_4^=$ is present in marine aerosol (Ovadnevaite et al., 2014). Therefore, the possibility that the HR-ToF-AMS may underestimate the nss-$SO_4^=$ concentration by missing some fraction of it, is highly unlikely. Conversely, sea-salt is considered a refractory component for the AMS, which means that sea-salt components tend to evaporate inefficiently within the AMS oven. This, together with the small contribution of sulfate in sea-salt (only 7.7% in mass) and considering the size distribution of sea-salt, that mostly falls outside the operative range of the AMS, makes the contribution of sea-salt-sulfate in AMS measurements usually negligible. In any case, in principle it may be possible to have an overestimation of the nss-$SO_4^=$ in case of high sea-salt-$SO_4^=$ contribution. Ovadnevaite et al. (2014) quantified these cases, concluding that a non-negligible contribution of sea-salt-$SO_4^=$ in the MHD database can be observed only for cases of low sulfate and high sea-salt concentrations associated to high wind speed events during winter months, when the contribution of sulfate is in any case close to the detection limit and negligible with respect to the high biological activity period. Similarly, Saliba et al. (2020), presenting the NAAMES database, states that non-refractory-$SO_4^=$ (measured by AMS) excludes refractory particles that likely contain the majority of sea-salt sulfate and that it is therefore approximately equivalent to nss-sulfate. To support this statement in a more quantitative way, we compared the concentrations of sea-salt and sulfate reported by Saliba et al. (2020), assuming a 7.7% $SO_4^=$ contribution in sea-salt: only during the winter cruise the contribution of sea-salt-$SO_4^=$ to the total AMS-$SO_4^=$ signal is non-negligible (54%), while in the other cruises it is around 10%. Anyway, this estimate is biased by the different cut-off of the samples used for sea-salt analysis (1.1 µm) and the AMS (~0.8 µm), which makes these numbers very likely overestimated.

Finally, if the reviewer is instead worried about the potential presence of nss-$SO_4^=$ in particles larger than the AMS upper cut-off, this may be true [and maybe even more for MSA (Rinaldi et al., 2011)] but we clearly state in the manuscript that our dataset refers to submicrometer particles, which falls in the size range of AMS and which are the more relevant climatically.

The authors use HYSPLIT driven by the Global Data Assimilation System (GDAS1) (1° × 1°) of the National Centers for Environmental Prediction (NCEP) to calculate back-trajectories (section 2.3). A different reanalysis, ERA5, is used to obtain meteorological predictor variables for machine learning methods (section 2.5), as well as the BLH used to analyze HYSPLIT-derived back trajectories (section 3.1.1). Can the use of different reanalyses in different parts of the study introduce inconsistencies?

In this study, the GDAS1 data set is only utilized to generate the Back-trajectories data, as one of the archived datasets in the HYSPLIT system. To achieve high spatial resolution (0.25° × 0.25°) of biogenic sulfur aerosol concentrations, we use the ERA5 dataset as predictors in machine learning models. The BLH was extracted along BTs in the same way as other atmospheric predictors because it also serves as a predictor. Indeed, all predictors included in model training were retrieved in the same way, therefore we do not believe such a strategy will create errors or uncertainties in the current study.

Section 3.3: please consider reporting other metrics, like the Prediction-Observation linear slope (which would be 1 for perfect model predictions) -- OK, this is shown in Fig. 6 and 7. Just consider introducing this metric in section 3.3.

We added the slope value as a metric to evaluate ML models in Figures 3 and 4 too. The following clause complements Section 3.3 in the revised manuscript.

"The predicted-observed linear slope is the last metric used to evaluate the performance of ML models. It determines the rate of change of the predicted variable concerning the observed variable and should be close to unity for skilled model predictions."

We modified panels (a-b) of Figure 6 and panels (a-d) of Figure 7 to make observations on the x-axis that are consistent with the explanation given above.

L272: How can this procedure prove causal relationships?

We agree with the reviewer that the sentence is misleading. We rephrased the sentence in the revised manuscript by eliminating the first part, now it reads:

"We used multilinear regression to assess the contribution of each predictor to MSA and nss-$SO_4^=$ variations."

Specific

L25, L85…: “constructed” >> “reconstructed”

Corrected.

L27: what is the “ensemble” ML method? OK, later defined as "regression ensemble"

We also added the word “regression ensemble” instead of “ensemble” in the specified line.

L42: marine phytoplankton >> marine microbes (phytoplankton are not the only DMS producers)

marine microbes replaced marine phytoplankton.

L49: elevated temperature and solar radiation >> elevated temperature OR solar radiation

Corrected.

L101 and paragraph: please revise whether the AMOC is the phenomenon you actually want to highlight here. Perhaps a mention to the Gulf Stream and the North Atlantic Current is enough (which indeed are components of the much wider phenomenon termed AMOC).

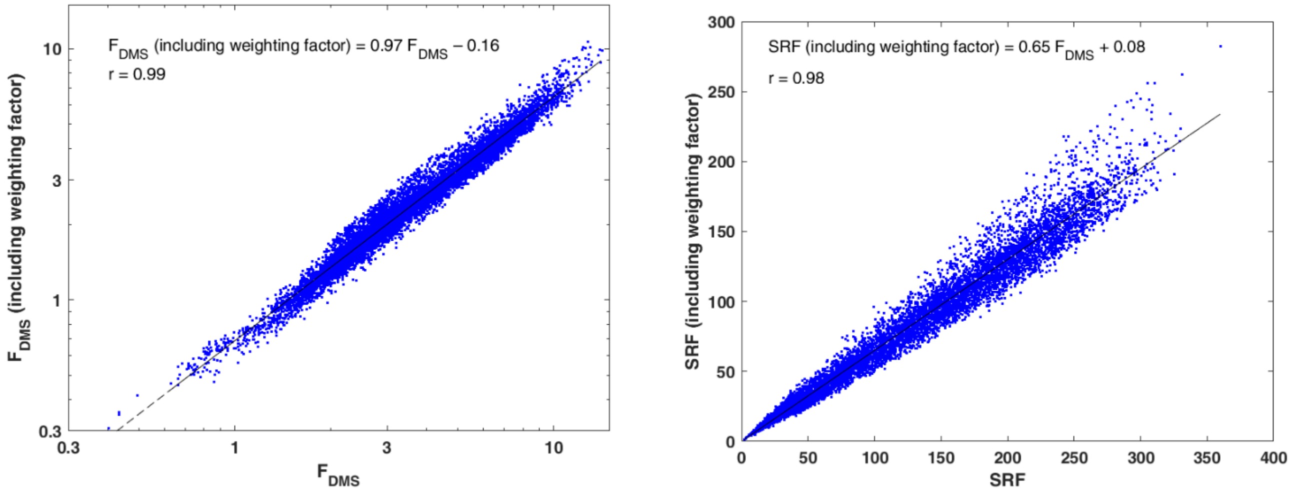
We agree to the reviewer's suggestion and highlighted the Gulf Stream instead of AMOC. The paragraph has been updated as follows:

“The key climate-relevant features in the study domain are the Gulf Strem, its northern extension towards Europe known as the North Atlantic Current (NAC), and the cyclonic subpolar gyre (SPG) (Rhein et al., 2011). The Gulf Stream is a warm Atlantic Ocean flow that begins in the Gulf of Mexico and moves through the Straits of Florida before continuing up the eastern coast of the United States (Buckley and Marshall, 2016). These warm northward-flowing waters meet the cold southward-flowing waters of the Labrador Current and the western boundary current of the cyclonic subpolar gyre, ultimately turning east and heading toward Northwest Europe as the NAC. The NAC then splits into multiple branches that enter the subpolar gyre, one of which passes via the Iceland Basin and the other through the Rockall trough (Fratantoni, 2001). The NA SPG extends from 45° N to around 65° N and comprises the sills between Greenland, Iceland, the Faroe Islands, and Scotland. Such circulation phenomena are crucial for the modulation of the temperate climate of north-western Europe (Marzocchi et al., 2015), and the dynamics of SPG determine the rate of deep and intermediate water formation (sinking dense and cold surface waters through air-sea heat exchanges in the wintertime) particularly in the Labrador Sea (Katsman et al., 2004).

Accordingly, they contribute to the regional changes of primary production and the subsequent biogenic emissions in the study domain.”

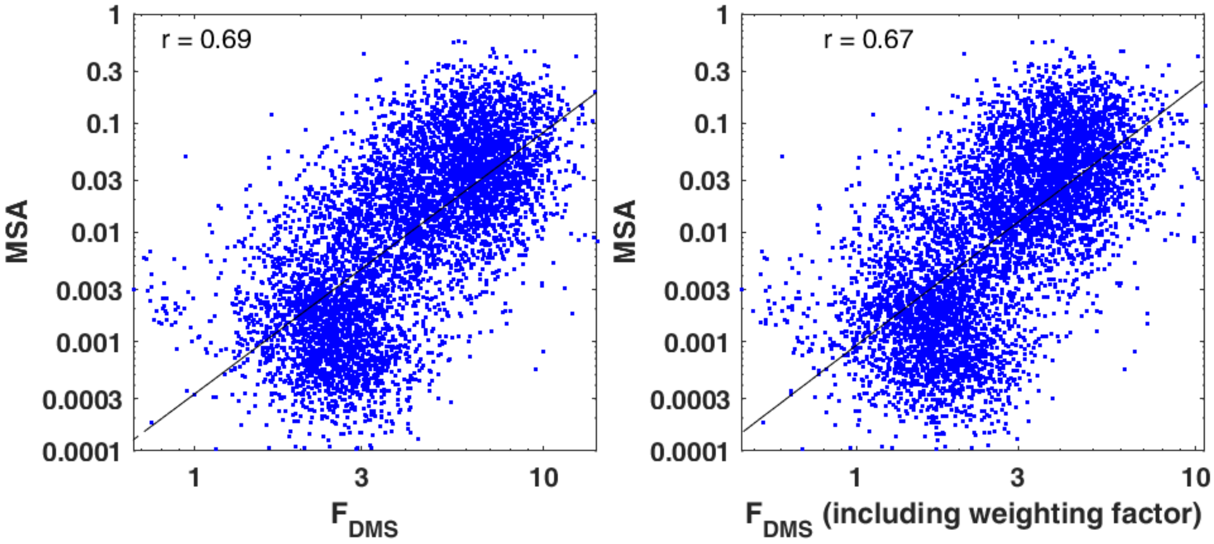
L238: were predictors averaged with or without the weighting factor e^(-t/72) used to compute R_0 and R_B? it would make sense to apply this weighting when using the meteorology along the BTs as predictor.

Thank you for pointing this out. We compared the weighted average F_{DMS} and SRF along BTs (as used in the present manuscript) to the same values when the weighting factor e^(-t/72) is included. The results reveal a strong connection between them ($r = 0.99$ for F_{DMS} and $r = 0.98$ for SRF), as shown in the following scatter plots.

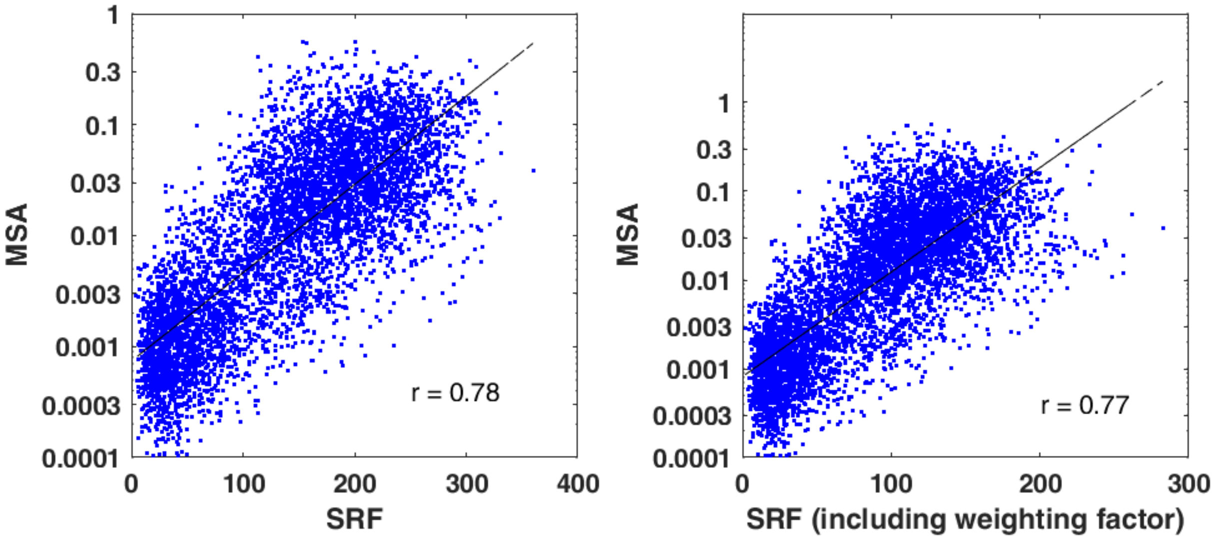


(Left) Comparison of F_{DMS} values with and without incorporating the weighting factor. (Right) Comparison of SRF values with and without incorporating the weighting factor.

In addition, we compared the F_{DMS}/ SRF (with and without the weighting factor) and MSA measurements at Mace Head. The results (figures below) show that incorporating the weighting factor does not change the relationship between predictors (*e.g.*, F_{DMS} & SRF) and response (*e.g.*, MSA). This may be due to the fact that the removal of submicron aerosol particles is negligible over a 1–3 days transport time. As a consequence, for this study, whether or not incorporating this weighting factor does not have a significant impact on the analysis results, we retained using the weighted mean along BTs without including the weighting factor.



Scatter plots between F_{DMS} at the selected marine air masses and the in-situ observed MSA concentrations at Mace Head.



Scatter plots between SRF at the selected marine air masses and the in-situ observed MSA concentrations at Mace Head.

L272: was MLR applied to untransformed or log-transformed data (as done for the correlation analysis)?

Yes, it was. We clarified this point in the caption of the multilinear regression table (Table 4; in the revised manuscript).

Typos

L232: “NAAMEAS” cruises

Corrected.

L477: “Quantitively”

Corrected.

L529: southern >> southward

Corrected.

References

Buckley MW, Marshall J. Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review. Reviews of Geophysics 2016; 54: 5-63.

Chen YL, Xu L, Humphry T, Hettiyadura APS, Ovadnevaite J, Huang S, et al. Response of the Aerodyne Aerosol Mass Spectrometer to Inorganic Sulfates and Organosulfur Compounds: Applications in Field and Laboratory Measurements. Environmental Science & Technology 2019; 53: 5176-5186.

DeCarlo PF, Kimmel JR, Trimborn A, Northway MJ, Jayne JT, Aiken AC, et al. Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer. Analytical Chemistry 2006; 78: 8281-8289.

Fratantoni DM. North Atlantic surface circulation during the 1990's observed with satellite-tracked drifters. Journal of Geophysical Research-Oceans 2001; 106: 22067-22093.

Katsman CA, Spall MA, Pickart RS. Boundary current eddies and their role in the restratification of the Labrador Sea. Journal of Physical Oceanography 2004; 34: 1967-1983.

Marzocchi A, Hirschi JJM, Holliday NP, Cunningham SA, Blaker AT, Coward AC. The North Atlantic subpolar circulation in an eddy-resolving global ocean model. Journal of Marine Systems 2015; 142: 126-143.

Ovadnevaite J, Ceburnis D, Leinert S, Dall'Osto M, Canagaratna M, O'Doherty S, et al. Submicron NE Atlantic marine aerosol chemical composition and abundance: Seasonal trends and air mass categorization. Journal of Geophysical Research-Atmospheres 2014; 119: 11850-11863.

Rhein M, Kieke D, Huttl-Kabus S, Roessler A, Mertens C, Meissner R, et al. Deep water formation, the subpolar gyre, and the meridional overturning circulation in the subpolar North Atlantic. Deep-Sea Research Part Ii-Topical Studies in Oceanography 2011; 58: 1819-1832.

Rinaldi M, Decesari S, Carbone C, Finessi E, Fuzzi S, Ceburnis D, et al. Evidence of a natural marine source of oxalic acid and a possible link to glyoxal. Journal of Geophysical Research-Atmospheres 2011; 116.

Saliba G, Chen CL, Lewis S, Russell LM, Quinn PK, Bates TS, et al. Seasonal Differences and Variability of Concentrations, Chemical Composition, and Cloud Condensation Nuclei of Marine Aerosol Over the North Atlantic. Journal of Geophysical Research-Atmospheres 2020; 125.