Answers to Reviewers

We would like to express our gratitude to the editor and the reviewers for their positive feedback on the manuscript. We revised the manuscript after considering the minor comments raised. In the responses below, the reviewer comments are in black text, and our responses are in blue and *the main text modifications to the revised manuscript are in italics*.

Reviewer 2

The paper has been significantly improved after incorporation of the reviewer's suggestions. My main concerns have been satisfactorily addressed. Therefore, I support publication after minor revisions. Below I list additional minor comments. I prompt the authors to incorporate them to further improve the manuscript:

1) The authors could mention the importance of reconstructing atmospheric MSA fields and its sources. MSA is widely used tracer of marine biological productivity. In the North Atlantic subpolar gyre, MSA from ice cores (deposited from the atmosphere) has been used to reconstruct decadal to centennial scale variability of marine productivity. An appropriate citation is Osman et al. 2019 Nature (https://www.nature.com/articles/s41586-019-1181-8).

We thank the reviewer for pointing out this valuable study. We modified the 1st paragraph in the introduction Section to read as:

"As a result, biogenic sulfur aerosols play a central role in ocean-atmosphere interactions and regional climate change, and it is critical to parameterize and characterize biogenic MSA and nss-$SO_4^=$ across different sea areas and identify their sources to constrain the past, current and future climate impacts of both species (Hodshire et al., 2019; Gondwe et al., 2003). For instance, MSA observations from Greenlandic ice cores have been used to study the variability of subarctic Atlantic Ocean productivity from decadal to centennial time scales (Osman et al., 2019)."

2) The inclusion of the model:obs slope metric is very welcome. This metric is typically around 0.80 on a log-log scale. Thus, even though the ML approaches used in this study significantly improve the prediction of nssSO4 and MSA in comparison to previous studies, they still tend to overestimate low

extremes and underestimate high extremes, as clearly seen in several figures (3, 4, 6 and 7). This is a common issue in statistical models. I prompt the authors to add a cautionary note on this source of uncertainty, which limits our quantitative understanding of extreme emission events.

We rephrased the 3rd paragraph in Section 4.1 to read as:

"Importantly, the implemented ML models can reconstruct MSA and nss-$SO_4^=$ daily time series characteristics with remarkable consistency between observed and predicted data, except for extremely high and low concentrations. This is mostly due to the low probability of such concentrations in the observed dataset, which inhibits ML models from reconstructing them. The quantitative comprehension of exceptional emission extremes is not addressed in this study; nonetheless, their occurrence and possible implications deserve to be investigated in future studies."

3) It appears that the term "error" should be replaced by "deviation", i.e., replace RMSE and MAE by RMSD and MAD, because observations have their own uncertainties. Strictly speaking, error should be used only when truth is known (.e.g in theoretical studies).

We agree with the reviewer on the "philosophical" point of view that observations have their uncertainties and do not necessarily represent the truth. On the other hand, in the manuscript we made use of the standard parameters and terminology used in a number of papers where model outputs are validated against observations. We prefer to keep these terms as they are (MAE and RMSE) to be consistent with the vast majority of the literature and to avoid ingenerating confusion in the readers about the main statistical terms. Indeed, the mean absolute deviation (MAD) and the root mean square deviation (RMSD) are used to measure the deviation or variability in a dataset from its mean value. For example, MAD is calculated as the absolute difference between each data point and the population mean, then, the average of these absolute differences is obtained. Indeed, these measures are not possible in our case.

4) Please, review thoroughly the references to pick potential typos or minor spelling errors.

Reviewed and checked.