

The reviewer comments are in black font and our responses are in blue.

### Reply to Wu Sun (reviewer #1)

In this work, Aydin et al. painstakingly recovered a 52,500-year record of atmospheric carbonyl sulfide from a South Pole ice core, with great care and attention devoted to correcting for the post-depositional COS production from sea salt aerosols and other artifacts during extraction. The resulting data set is a valuable contribution to the atmospheric history of carbonyl sulfide and will enable the climate and ecosystem modeling community to better understand biospheric changes since the last ice age.

Despite its scientific significance, the organization of the manuscript may hinder its key findings from being grasped by a broad audience. For this reason, I suggest clarifying the methods, especially the rationale behind every correction and sensitivity test, and streamlining the presentation of the main messages to strengthen the work. Below I list a few high-level issues followed by specific line-by-line comments.

My issues with the methods are:

- For the calculation of  $ssNa^+$  (Sect. 2.3), please check the denominator in the right hand side of Eq. (6). My derivation says it should be " $1 - R_m/R_t$ " in the denominator (see attached slides). In other words, there is no reason for Eqs. (6) and (7) to have different denominators.

The reviewer is correct. Both denominators should display a subtraction of  $R_m/R_t$  from 1. This is now corrected. This change results in a linear 4% increase in the calculated  $ssNa$  but no change in the data analysis.

- For the regression between COS and  $ssNa^+$  (Sect. 2.4), it was not clear why errors in the response variable need to be scaled with a multiplier ( $\alpha$  in Eq. 8). This multiplier did not appear in Eq. (5).

We included  $\alpha$  as a multiplicative scaler in eq. 8 to demonstrate the impact of a possible bias in determination of the measurement errors (e.g. consistent underestimation of the measurement errors) on the regression results. The  $\alpha$  parameter is 1 for three of the scenarios (G1, G2, H1) and its distribution is determined by the Bayesian algorithm in the other scenarios, allowing us to demonstrate the difference between using calculated measurement errors versus what the impact on the results is if the measurement errors are biased low. We included new text in section 2.4 to clarify (P6, L195-196): "Inclusion of the  $\alpha$  parameter allows quantification of the sensitivity of the correction slope to possible bias of COS measurement errors."

- It was also unclear how measurement errors in COS and  $ssNa^+$  (Sect. 2.5) were incorporated into the Bayesian errors-in-variables regression between COS and  $ssNa^+$  (Sect. 2.4).

The COS errors are directly used as measurement errors (i.e. errors in  $y$ , denoted  $y_{err}$  in equations 8, 9, and 10). The  $ssNa$  errors are used as  $x_{err}$  in eq. 4. We added text to section 2.4 to clarify (P7 L201-202): "The error estimates for the COS and  $ssNa$  measurements are directly incorporated in the errors-in-variables regressions as they substitute for  $y_{err}$  and  $x_{err}$  in equations

4 and 8, allowing us to propagate the uncertainty in the measurements to the slope of the relationship between ssNa-COS.”

- Different scenarios for correcting for ssNa<sup>+</sup>-produced excess COS (Table 1) should be clearly outlined in the methods section, with the rationale explained. It was not until I read the results did I get a hint of why these scenarios were needed and why they were set up like these. The choice of the averaging window size appears ad hoc.

The manuscript is revised to introduce the different scenarios under methods in section 2.4 with a sufficient amount of detail about their purpose (P6 180-189). Some details of the scenarios are still presented under results in section 3.1 because decoupling some of the descriptions from the results of the scenarios makes the text incomprehensible.

The reviewer is correct that the exact values of the averaging windows are ad hoc. The 100-y averaging window is chosen to be higher than the span of the ice age of a typical sample (about 10 y during the glacial period) and yield a higher statistical significance. The 1100-y window is chosen to match the averaging window of the smoothing applied to the COS record to diminish the influence of the spikes on the regressions. The results are robust with respect to the averaging windows despite the large range explored in data averaging.

- Too many scenarios overlaid on the same plot (e.g., Figs. 4b and 5) make it hard to discern the key information. Consider presenting only scenarios that are most robust and relevant to the interpretations.

The results from all scenarios are displayed together in Fig. 4b because one of the points we are trying to get across is the fact that the interpretations rely only on the features that are common to all scenarios: low during the glacial period, lowest during the LGM, and high during the Holocene. In the original version of the manuscript, Fig. 5 included two less scenarios than Fig. 4b. In the revised version, we eliminated three other scenarios (two from glacial and one from the Holocene) from Fig. 5, reducing the number of scenarios displayed in this figure to three. Fig. 5 is now a fairly plain figure.

The results section needs to provide answers to questions raised in the introduction. Currently it is a mix of methods, results, technical arguments, and discussion, making it challenging to navigate. The main issues are:

- Many paragraphs describe the nitty-gritty details of various corrections and make reference to figures in the appendices. Important as they are to ensuring data quality, these are probably not the high-level findings you want the readers to walk away with. Consider moving them to methods or the appendices.
- Descriptions of the correction for ssNa<sup>+</sup>-produced excess COS are scattered throughout the results section, making it difficult to follow. Consider consolidating the main points about this under a single subsection and offload nonessential details to supplementary materials.

Based on the reviewer comments, what used to be sections 3.2, 3.3, and 3.4 are now moved to the appendices. These sections included many of the additional data analyses including sensitivity tests for G1 and G4 scenarios, testing the validity of the glacial same-age correlation as a climate signal, and statistical tests regarding the COS rise concurrent with the deglaciation. Atmospheric COS variability inferred from different scenarios are still presented under results in

section 3.1 with a modified title “Results of different analysis scenarios: Inferred atmospheric COS variability after the ssNa<sup>+</sup> correction.” Section 3.1 contains the details of exactly what was done under each analysis scenario.

- It was not clear how the "climate-driven" correlation between ssNa<sup>+</sup> and COS was disentangled from the production artifact. Shouldn't the ssNa<sup>+</sup>COS production correction be applied first before one can make any robust inference of a climate-driven relationship?

By definition, the climate relationship is observed over time (same-age) and the production relationship is observed over same-depth. If the delta-age is large enough, which is the case for the South Pole ice core, separate same-depth and same-age regressions can reveal both the production and climate driven signal as we show in the manuscript with the independent same-depth and same-age analyses for G1 and H1 scenarios. In a typical data analysis approach, untangling the two relationships from each other would commonly be accomplished by correcting for the production first, followed by quantifying and accounting for the same-age relationship, then cycling back to the same-depth correction, and so forth until the successive corrections do not statistically differ from each other. The simultaneous correction algorithm we deploy here in G4 and H2 scenarios allows carrying out this operation in one step within the Bayesian framework. Note that the software used in this analysis is publicly available and the code is made available as supplemental information. The unique aspect of the code used in this analysis is that it does not make use of an iterative process to accomplish this task.

- West Antarctica and Taylor Dome ice cores (L233–243) need to be described in the methods.

We realize that the names of the ice cores (i.e. WAIS Divide and Taylor Dome) were not explicitly mentioned in the introduction paragraph, likely obscuring the fact that the measurements from these ice cores were presented in prior publications. In the revised version, we modified the introduction paragraph to refer to these ice cores by their names (L50-51) where we describe the findings from these two ice cores, citing the relevant references (L48-55). The references are also provided in the results section where we refer to the results from these ice cores (P8 L263-266) and then again in section 3.2 where the records from different cores are compared with each other (P12 L431). The measurements from these ice cores are published in archival journals including all relevant methods.

With respect to the interpretations of the data, there are a few issues:

- The measurements presented here only show ssNa<sup>+</sup> but say nothing about what the anions in sea salts are. Thus, it is unclear whether sea salt aerosols contain *organic* sulfur compounds and the analogy between COS dark production in the ocean and that in ice (L462–L476) does not seem to be empirically supported. Experimental evidence suggests that COS is produced from organic sulfur compounds (Modiri Gharehveran and Shah, 2018, <https://doi.org/10.1021/acs.est.8b01618>) such as cysteine, methionine, and chromophoric dissolved organic matter (CDOM). In any case, a reduced sulfur precursor is needed to produce COS in the absence of sulfate-reducing microbes.

The most abundant anion in the ice core is Cl(-), which is strongly correlated with Na(+) since Cl(-) is sourced primarily from NaCl. Sulfate is another important anion but is sourced primarily

from inorganic sulfur (SO<sub>2</sub>) with an additional volcanic component and experiences some post-depositional processing. We conducted regression analyses between COS and all major ions measured in the South Pole ice core, including Cl(-) and SO<sub>4</sub>(2-). The Cl(-) regression results display linear relationships with COS much like Na(+) for the reason stated above. There is no relationship between COS and SO<sub>4</sub>(2-). We do not include these other regression analyses in the manuscript because they do not provide additional insight. There are no measurements of organic sulfur from the South Pole ice core that can be used in our analyses, and as far as we know, there are no commonly accepted proxies for organic sulfur content in ice cores.

We are aware that an organic sulfur source is needed for abiotic COS production, e.g. (L619-621): “The commonly postulated abiotic process involves reactions between carbonyl groups and thiyl radicals derived from organic sulfur (Flöck et al., 1997; Modiri Gharehveran and Shah, 2018; Lennartz et al., 2017; 2019; Pos et al., 1998; Zepp and Andreae, 1994).” We do not make any claims of full empirical support for the proposed mechanism, e.g (L605-606): “In the absence of auxiliary data, we can only speculate on viable mechanisms for the COS production in the firn.” To clarify, this sentence has been revised to read “In the absence of direct empirical evidence, we can only speculate on viable mechanisms for the COS production in the firn.” We also added a sentence to this paragraph that explicitly states the need for an organic sulfur source for abiotic COS production before the already existing sentence that suggests marine aerosols could serve as this source (L616-618): “Abiotic COS production also requires an organic sulfur source. Antarctica is far from continental land masses, suggesting marine aerosols could also be the primary source of organic sulfur.”

- Given the large uncertainty in the dimethyl sulfide (DMS) contribution to the global ocean COS budget (53 to 680 GgS yr<sup>-1</sup>; Jernigan et al., 2022, <https://doi.org/10.1029/2021GL096838>), it seems hasty to suggest that atmospheric COS is sensitive to changes in DMS emissions.

The manuscript does not propose that the paleoatmospheric variability of COS reflects changes in DMS emissions, although this is certainly possible within the uncertainties. The sentence on L729-731 makes this clearer: “We cannot quantify how much of the ocean COS source increase results from DMS versus COS and CS<sub>2</sub> because of the complexities of ocean production mechanisms of sulfur gases and the uncertainties in their contribution to the atmospheric COS budget.” Instead, we argue that a large increase in ocean COS emissions is unlikely to happen without an increase in ocean DMS emissions as stated on L735-736). This reviewer comment seems related to the first specific comment below. We added some additional discussion to clarify and strengthen the argument about inferences related to DMS (L736-740) that are also discussed below.

- Without paleoceanographic evidence, it seems speculative to single out changes in coastal upwelling zones and low-latitude oceans as likely contributors to the deglacial rise in COS. At best this represents one scenario among many possibilities.

Inline with the reviewer’s suggestion, we merely suggest coastal upwelling zones are one of the possible contributors to deglacial rise in COS. We discuss multiple possible mechanisms and the relevant sentence (L726-728) comes at the end of the discussion.

- It was not clear why a climate-driven relationship between  $\text{ssNa}^+$  and COS has to be an anticorrelation. What are the biogeochemical or climatological reasons behind this?

There is no a priori assumption that the climate-driven relationship has to be an anticorrelation. The data analyses reveal an anticorrelation during the glacial period and a positive one during the Holocene. We opted not to speculate on the possible causes in this manuscript because the causes of  $\text{ssNa}$  variability in the glacial sections of the South Pole ice core have not yet made it into the literature.

### Specific comments

- L24–25: Seems speculative. The results may allow us to infer a likely increase in total COS emissions, but not in the emissions of each precursor.

This is an interesting point. From our perspective, the least speculative option is to suggest that emissions of all three gases increased. Of course, the details of where exactly most of the emission changes can come from and how much from each gas requires a detailed modeling effort, beyond the scope of this manuscript. We added a sentence to this effect on L739-740 in the revised version.

All ocean sulfur gas emissions ultimately stem from organic life in the ocean. Even though net emissions of COS and  $\text{CS}_2$  are geographically decoupled, DMS emissions overlap with net emissions of both gases, and all three gases are emitted at coastal upwelling zones. Additionally, when all three gases increase, relatively smaller increases are required for each gas to achieve higher overall OCS in the atmosphere. If only COS and DMS or  $\text{CS}_2$  and DMS increased, due to changes in upwelling regimes only in the high latitudes or only in the low latitudes, for example, much larger changes in gas emissions would be required within that region to make up for the missing increase from the other region.

There is paleoclimate evidence that upwelling regimes in both high and low latitudes were different during the LGM, and sea level rise during the deglaciation (100-120 m) had dramatic impact on coastal regions and processes close to the continental shelves. We provide multiple references in the manuscript in the relevant paragraph (L720-735) which we supported with a few more in the revised version.

- L33: "COS and  $\text{CS}_2$  are produced primarily by photochemical reactions" - COS and  $\text{CS}_2$  are also produced in the dark. See Lennartz et al. (2019) *Ocean Sci.* (<https://doi.org/10.5194/os-15-1071-2019>) and Modiri Gharehveran and Shah (2018) *ES&T* (<https://doi.org/10.1021/acs.est.8b01618>).

The evidence for dark production of  $\text{CS}_2$  in the ocean is limited but we modified the text to reflect this possible production pathway for  $\text{CS}_2$ .

- L35: "Warmer waters can act as a seasonal sink due to temperature dependent loss to hydrolysis" - But both COS and  $\text{CS}_2$  are less soluble in warmer waters (De Bruyn et al., 1995, <https://doi.org/10.1029/95JD00217>), and wouldn't this lead to more outgassing from the ocean?

This is addressed in Appendix E (previously Appendix B). Briefly, the hydrolysis loss rate of COS is more sensitive to a unit temperature change than its solubility. Observations leave little room for doubt that COS emissions are low and  $\text{CS}_2$  emissions are high at low latitudes (Lennartz et al., *Earth Syst. Sci. Data*, 12, 591–609, 2020). We added this citation to the manuscript.

- L52: It would be helpful to add a brief note on how the bubble–clathrate transition zone (BCTZ) affects the preservation of ancient air in ice cores for those who are not familiar with the BCTZ.

Dry extraction measurements from the BCTZ display large negative biases (L52 and L82-83). The readers are referred to a citation provided in the text (Aydin et al., JGR, 2016) for more detailed information.

- L80: How were the extraction efficiencies of the wet and dry methods characterized? If these methods have been previously examined, a citation would help.

Wet extraction efficiency is close to 100% because solubility of air in water is negligibly small in the context of determining extraction efficiency; Nicewonger et al. (2020) estimate that about 0.7% of the air is left dissolved in the melt water in our extraction vessels. The dry extraction efficiency can be estimated from comparison of total air content data, which is estimated reasonably well by wet extraction measurements from the same core like what was done here, or by comparisons with total air content data from the same core which may be available from a different lab previously by Aydin et al. (2016). We added citations to the text.

- L112, Eq. (1): Is  $\alpha$  the ratio of aqueous concentration divided by gaseous concentration?

Yes. It is similar to the dimensionless Henry's Law solubility except it is the value of effective solubility instead of saturation value. We added this information to the manuscript and changed this coefficient to  $h$  to avoid confusion ().

- L165, Eq. (8): Is  $\alpha$  here the same as that in Eq. (1)? If not, please use a different symbol.

We changed Eq. (1) and kept Eq. (8) the same.

- L191–192: The numbers of effective sample sizes should be reported properly as 4900, 2500, etc.

Done.

- L197: "The  $2\sigma$  uncertainty ranges shown in the figures represent" - Which figures?

Figs. 4b and 5 (L221).

- L215–L292: This section seems to belong to the supplementary material or methods. I would distill a few key messages only to put in the results section.

For a different audience, this information is necessary to include here.

- L215: "COS was measured at the over the length of ..." - Check typos here.

Done.

- L220: Fig. A1a shows the amount of gas extracted, not gas extraction efficiency per se.

We modified the sentence to reflect this nuance (L242).

- L222–232: This paragraph left me wondering: are the spikes real or not?

We cannot say with confidence and state “it seems unlikely that the COS mixing ratio in the glacial atmosphere varied abruptly at the magnitude and frequency of these spikes (L229-230).”

- L255: "This reversal occurs at a depth where ice impurity concentrations are increasing steeply (Fig. 1b)." - I believe here you intended to reference Fig. 1c.

Yes, corrected.

- L256–264: This paragraph seems to belong to the methods.

This paragraph is necessary to understand the following paragraph.

- L265: "significant" -> "statistically significant"?

Thanks, we changed the sentence to read “statistically significant.”

- L266–267: "The slope is stronger ... with ssNa" - If the mechanism by which COS is produced from non-sea-salt aerosols differs from sea salt aerosol-caused COS production, can the slopes be compared on the same scale?

The slopes cannot but the significance can. We changed the sentence to reflect this (L289).

- L269–271: Skimming these sentences the first time, I was confused how an  $R^2$  value of 0.04 could serve as the basis for the correction. The next time I realized that this was not the correlation between excess COS (the noise, which is unknown) and  $ssNa^+$ , but that between the total COS (signal + noise) and  $ssNa^+$ . You might want to add a brief note somewhere to get this point across more effectively.

Thanks, we modified the sentence to include this information (L287-289).

- L276: "In fact, the presence of spikes is a contributing factor to the low  $R^2$  of the correlations despite the high significance" - You might want to point the readers to scenario G2 here to show that the issue of spikes has been taken into account.

Good idea. Done (L298).

- L293–304: This seems to belong to the methods.

This paragraph has been moved to the methods with minor modifications (L171-183).

- L303–304: It was not obvious to me how these analysis scenarios supported the idea that "the relationship between ssNa and COS is driven primarily by millennial scale variability."

We show higher frequency variability do not correlate and lower frequencies do as demonstrated by higher significance of correlation for more smoothed records.

- L317–318: Why was it necessary to scale up the errors? Section 2.4 asks me to go to section 3.1, but section 3.1 refers me back to section 2.4.

This is mostly done out of abundance of caution. The errors are scaled up to test the sensitivity of results to possible underestimation (negative bias) of errors. We included a sentence to this effect on L189-190.

- L334–342: Why are measurements of the glacial period and the Holocene corrected separately? Do you expect the relationship between  $ssNa^+$  and excess COS to differ between the last glacial period and the Holocene?

There are reasons they could be different because during the deglaciation  $ssNa$  sources regions and the concentration of whatever else is co-deposited can change as well as accumulation rate. In the end, the results show a minor difference, implying a correction to the entire record using only the glacial slope does not change the outcome significantly enough to impact the interpretation (Fig. 4b).

- L345: "150 ppt" and "90% lower" - Check the numbers. It's not like that the Holocene has a COS level at 1500 ppt. Also the referenced figure does not seem to tell this information.

Thanks for catching this. The sentence was initially written in a reverse sense (i.e. 90% rise from the glacial period). We revised the sentence to refer to the ppt difference instead (L385).

- L370–376: The point of these three analyses is lost on me. How do the regressions of smoothed  $ssNa^+$  onto unsmoothed  $ssNa^+$  support the validity of climate-driven same-age anticorrelation between  $ssNa^+$  and COS?

The key component of the analysis is that, in all three control scenarios, one of  $ssNa$  records is on the gas chronology while the other  $ssNa$  record is on ice chronology. We realize this may be confusing for readers who are not ice core scientists, but this is a valid way to test whether the applied same-depth correction results in, or how much it contributes to, the same-age correlation. These tests are akin to testing the autocorrelation of the  $ssNa$  record at a given lag, but instead of a fixed lag, the delta-age for the ice core is used. We modified the associated text in the manuscript to better explain what was done and the meaning of it (L1106-1112). Note that this entire section has been moved to Appendix C in the revised manuscript.

- L454: "COS production in the firn is approximated by an advective-diffusive model of the South Pole firn" - Has this correction been applied to the present study?

The model is presented as a proof of concept. No correction has been made to the measurements using the firn model. We reworded this sentence and added additional clarifying text in the revised version (L610-614).

- L487: "It is possible atmospheric COS is sensitive to changes in ocean DMS emissions modulated by winter sea ice" - According to Lana et al. (2011) climatology, DMS emission hotspots seem to lie beyond the sea-ice covered regions of the Southern Ocean.

Seasonal sea ice is different than the climatological means used in DMS inventories in that dissipation of winter sea ice commonly leads to enhanced DMS emissions, although the reasons remain an open scientific question. Two citations were included in the manuscript (Curran and Jones, 2000; King et al., 2019 L640). There are others suggesting different mechanisms, including one cited by Lana et al. (2011): Trevena, A. J., and G. B. Jones (2006), Dimethylsulphide and dimethylsulphoniopropionate in Antarctic sea ice and their release during sea ice melting, *Mar. Chem.*, 98(2–4), 210–222, doi:10.1016/j.marchem.2005.09.005.



- Fig. 1: For wet extraction results, maybe show only the solubility-corrected values (red line in 1a)? Captions seem misplaced for panels **b** and **c**.

The motivation here is to present the measured wet extraction results so they can see for themselves that the solubility correction itself does not introduce trends to the record.

- Fig. 2: I don't see the "black error bars" described in the caption.

Thanks for catching this. Black error bars were used in a previous version of this figure. Corrected the caption to refer to magenta circles.

- Fig. 4a: Isn't this panel a repeat of Fig. 1a, but with error bars? It may be removed because it adds little new information.

One of the purposes of Fig. 1a is to display the impact of the solubility correction. We opted not to use error bars in that figure for purposes of clarity. An errorbar plot version of the measurements are shown again in Fig4a for two reasons: 1) It provides an immediate comparison with the corrected records under different scenarios shown in Fig. 4b, 2) it displays the errorbars used in the Bayesian analysis that underpin the 2 sigma uncertainty bands of the corrected records in Fig. 4b.

- Fig. 5: Could be combined with Fig. 4 for the comparison.

This is a good idea for comparison purposes, but we suspect the figure will get too large with four panels and the accompanying caption to fit in one page.

- Fig. 6: Seems like a supplementary figure to me.

We anticipate this to be a compelling figure for ice core scientists, specifically for firn air modelers, who are interested in how gas production in the firn can influence ice core gas records.

We thank Dr Wu Sun for his insightful and detailed review of the manuscript and hope that our responses are sufficient to make the best use of his effort.