Response to Reviewer 1

We thank you for your in depth review of our paper. We have clarified the points you have raised with respect to both the model structure and language. Please see our detailed responses below.

Title and sub-titles:

- *The 2nd part of the title is not clearly supported by the results and should be changed to better reflect the paper content.*

We have updated our title to be more specific in line with the Reviewers comments.

'The ACCESS-AM2 climate model underestimates aerosol concentration in the Southern Ocean; improving aerosol representation could be problematic for the global energy balance'

- *Sub-title of section 3.4.2 is incorrect. The content describes data processing rather than statistical methods.*

We have updated this to 'Data Processing' as suggested

Vague language:

- *Language is imprecise in parts of the article, leaving the reader to guess the author's intended meaning. The discussion section stands out as particularly vague. One key example: in the first paragraph of section 6, the phrase 'as a whole' is unexplained, yet this seems to be the primary recommendation of the paper. The authors need to take the time and space to frame their hypothesised model development framework in more detail and with greater clarity to be convincing.*

We have made every effort to be more precise in our language. We have now revised the discussion, removing or explaining imprecise language, such as 'as a whole'. We have also more clearly laid out our argument and hope that the Reviewer may now be convinced as to its robustness.

For example:

Line 589: 'It also demonstrates the internal complexity of the aerosol population and the need to consider how each component influences the size and composition of the entire burden, rather than as individual (compositional) populations.'

We have also made our primary findings of the paper more clear in the conclusions:
Line 698:'We draw two main conclusions from this work, corresponding to two different regions of the Southern Ocean. We suggest that better capturing the biological influence on aerosol may lead to limited improvements in the aerosol-cloud-radiative system of the Southern Ocean's sea-ice regions, where the radiative bias is at its worst. We also show that in order to reduce the uncertainty of cloud feedbacks and the energy balance in the northern parts of the Southern Ocean, improving the aerosol alone is not effective (in fact is detrimental) but may be a pre-requisite for improving aerosol-cloud interactions due to the influence of aerosol composition and size.'

Other examples of vagueness in language that need to be addressed include:

- *In section 3.1, the description of the model indicates the ACCESS model was run in atmosphere-only mode. So, the model is essentially UM10.6 GA7.1 with GLOMAP-mode, using CMIP6 and CEDS emissions, nudged towards ERA5 data. The coupled aspect of the model seems irrelvant yet ACCESS is framed as the model being evaluated. What is ACCESS actually adding to the simulations?*

The Reviewer is correct that the model, run in atmosphere only mode, is based on UM10.6, GA7.1, however it also includes the Australian land-surface model CABLE, as opposed to the JULES model. This is stated on Line 130. The model, even when being run in atmosphere only mode, is referred to as ACCESS, with the second part of the name (eg. AM2) referring to the specific configuration. We make reference to the coupled model, as much of the documentation for the atmosphere only version is also captured by the documentation for the coupled model (eg. Bi et al. 2020).

- *Line 198 to 202: This meaning of this paragraph is hard to unravel. Clearer language is needed.*

We have revised this paragraph as follows to bring more clarity:

Line 221: 'We have also produced a daily, annually varying DMS dataset derived from output of the ocean component of ACCESS, ACCESS-OM2. We refer to this experiment as 'OM2 DMS'. Details of ACCESS-OM2 and the simulation used to produce the DMS output can be found in Kiss et al. (2020) and Sections 2.1 and 3.1 of Hayashida et al. (2021). ACCESS-OM2, in this case has used the atmospheric boundary conditions from the Control* experiment to drive the model (instead of reanalysis). We highlight that this DMS data set differs from the previous in that it is not a climatology - it is an annually varying dataset at a daily scale, able to respond to atmospheric and oceanic forcings. '

- *Line 304: GLOMAP-mode provides these values, though they may not have been selected by the authors. This description needs to clearly state this was a choice, rather than a model deficiency.*

It was not our intention to imply that this was a model deficiency. In our configuration, GLOMAP-mode outputs the N3, CCN50 and CCN70 by default, not CCN at a specific supersaturation. To more accurately represent the critical activation diameter for the Southern Ocean, we needed to calculate additional cut-offs, including CCN40, which are used in this work.

- *Figure 3 caption: 25th and 75th percentiles of what?*

The caption (Figure 3, 4 and 5) have now been updated to be more specific: for example:

Figure 3: '... For all subfigures, the 25th and 75th percentiles of the daily mean N10 are shown by the shaded range for the observations and control run. ...'

- *Line 340: Standard deviation of what? Calculated from which data?*

We have updated this to:

Line 383: 'The standard deviation of the daily mean N10 is also underestimated on average, where the control run for KGC has a mean standard deviation of $123\,\mathrm{cm}^{-3}$ compared to $342\,\mathrm{cm}^{-3}$ in the observations.'

We have also updated other sections of the text where it may be unclear as to which data we are referring to.

- *Line 463: The meaning of the final sentence is obscure and unreferenced.*

We have removed this sentence.

- *Constraint/constrained is used incorrectly in the article. I think the authors mean 'restricted to', 'in', or 'limited to'. Constraint has a specific meaning related to model uncertainty.*

We have updated two instances in the text in line with your suggestion.

- *These is extensive use of accronyms, which may be considered appropriate for some readers, but reduces readability. Particularly, readability is reduced by using accronyms for observation stations.*

We have removed reference to the SYO acronym (left in by mistake). We have also removed the acronym for Macquarie Island (MI). The acronyms for Kennaook/Cape Grim and ship-based campaigns have been retained.

Presentation of results with more obvious scientific reasonining:

Much of the manuscript needs to be rewritten to highlight key discoveries to the reader. The authors should consider where meaning is assumed and could be clarified. Additionally, the text often contains only statements about model-to-observation comparisons, without interpretation of meaning. Occassionally, statements conflict with results, which suggests they've not been considered deeply.

We have revised the Results section to be more clear and remove any statements of conflict. We note that we have a dedicated discussion section for more in depth interpretation of the results, which considers all of the results together rather than the individual as presented in Section 4. In line with a later comment by this Reviewer, we have tried to reduce the 'long-winded'-ness of this section but will keep the in-depth interpretation of meaning in the discussion section.

Some essential changes include:

- *Section 3.4. The first paragraph here is unneccessary. Nothing of value is added, so this should be removed.*

It has been removed

- *Line 335: First sentence is confusing.*

We have changed it to:

Line 379: 'Figure 3a-c shows the modelled and observed N10 concentration seasonal cycle for KCG, Macquarie Island and Syowa.'

- *Paragraph starting line 360: In this location, the NPF scheme test is the only sensitivity test to shift the model from biased low to biased high. Some interpretation of results is needed here. This result implies the persistent model bias might be partially overcome by implementing a more sophisticated NPF parametrization. Also, the simplicity of the NPF parametrization needs to be mentioned in this section to help the reader understand why the improvements are spatially restricted.*

We have included some discussion about the NPF results section. The Review raises a valid point that a more sophisticated NPF scheme may yield better results that can increase the small sized aerosol in a more realistic way. We have revised this paragraph to include the following:

Line 410: 'This large increase in small-sized aerosol may be a result of several factors, including the relative simplicity of the GLOMAP-mode BL NPF scheme (a binary scheme outlined in Section 3.2.1), the influence of terrestrial airmasses (which contain emissions of VOCs that mediate the BL NPF, despite our efforts to filter these influences out) or aerosol pre-cursors. A more complex NPF scheme, such as those discussed in the Introduction may yield more realistic results, while greater investigation into the observed and modelled aerosol and aerosol precursors is called for in the region.'

- *Section 4 has many long-winded descriptions that do not lead to insights or statements of how the results affect model interpretation.*

We have revised this Section to make it more concise. We note that much of our interpretation of the results is presented in Section 6.

- *Section 4.3: There is no mention of the BL NPF sensitivity test here, even though it is the only test to reduce activation ratios.*

Yes, good point. This was an attempt at being concise as we do not believe the BL NPF simulation to be realistic. But we can see that these results still deserve a mention as this Reviewer suggests. We have added the following:

Line 539: 'The BL NPF simulation in most cases, particularly for the northern regions, reduces the activation ratio, demonstrating its large production of small sized aerosol, which we suggest is unrealistic.
'

- *Line 563: The meaning of the first sentence is incongruous with the results.*

We have clarified this sentence to read:

Line 611: 'We find that BL NPF had little impact on the regions of the Southern Ocean and Antarctic least influenced by terrestrial airmasses of the mid-latitudes (south of 45S). For regions where terrestrial airmasses are common (eg. northern latitudes and KCG), turning on BL NPF strongly overestimates

small sized aerosol, which we suggest to be unrealistic. '

Missing detail and context:

As mentioned above, some sections are heavy with text, whilst others lack detail and critical information.

- *For example, where the Humphries data set is introduced, no context is provided for why it might be better, or more useful, than previous data sets. Furthermore, some sense of the motivation for including the specific sensitivity tests chosen would be extremely useful in the first paragraph of section 3.2.*

We have added some text around why the Humphries data set is used:

Line 247: 'Most of these observations have been described, collated, quality controlled, harmonised and evaluated in Humphries et al. (2023). The Humphries et al. (2023) paper provides the first seasonal and latitudinal description of Southern Ocean aerosol properties, providing an ideal basis from which to perform a modelling evaluation for this region.'

Additionally, we have now included some motivation for the sensitivity tests, as suggested by the Reviewer at the start of Section 3.2:

Line 170: 'These sensitivity tests range from realistic and established updates through to some experimental only changes. They include tests that bring the model in line with recent UM configurations (e.g. the inclusion of primary marine organics), the use of updated ancillary data (i.e. the new DMS climatology), and examining the applicability of existing parameterisations that are usually not used in this region (i.e. boundary layer new particle formation). Furthermore, changing the sea salt parameterisation and using a model-derived daily updating DMS field are more experimental but are useful for future model development.'

Other specific examples of missing detail/context include:

- *Line 207: Why are time-varying DMS datasets preferable? Need to say what value is added.*

We have included this information as suggested:

Line 228: 'The benefit of a daily varying dataset is that it is able to respond to atmospheric and oceanic forcings, such as sea surface temperatures or wind speed. This method can present, potentially, a more tightly coupled system, and if the parameterisation is accurate, yield more realistic DMS fields (including DMS in the water and the atmosphere). Some modelling groups are already adopting online DMS production (e.g. Bock et al. 2021), so this is a first step towards this goal for ACCESS. '

- *Line 222: 'underway' needs a description*

We have defined 'underway' as:

Line 256: ' 'underway' (automatic observations taken continuously while the ship is operating)'

- *Line 317: Why isn't the assumption made that the gridbox containing the observation would be the best comparison to make? There is no explaination for why the authors are even considering using a gridbox the the SW of the station.*

We have added a more detailed explanation of our choices re: grid box location.

Line 254: 'This is also true for KCG, where choosing a gridbox to the south-west of the station, as is normal practice for this location when studying baseline airmasses (which are not influenced by terrestrial air), resulted in poorer performance.'

- *Line379: The authors state 'This is a key area of development for GLOMAP-mode'. This statement needs to be put in context. The suggested model developments are only important if the priority is a model with increased skill at simulating aerosol concentrations over remote polar regions. The authors have assumed this is the case, with an implied further assumption that aerosol in these regions are more climatically important than aerosol elsewhere.*

This was not our intention. But to avoid confusion, we have removed this sentence.

- *Section 4.1: Some brief description of the overall under-prediction of CCN concentration and seasonal cycle amplitude, and what this implies should be given up front. Currently, this message is hidden amongst discussion of individual simulations.*

We have added the following:
Line 375: 'In this section we show that the ACCESS-AM2 model strongly underestimates both the N10 and CCN concentrations, which points to issues with the model's ability to accurately represent the aerosol population. We have limited our discussion in this section as the results will be analysed and interpreted together in Section 5. '

Old model version:

The authors have evaluated the impact of structural changes to model parametrizations, using a relatively old version of the GLOMAP-mode aerosol scheme, without reference to published model changes that would affect results. Results in this article need to be discussed with reference to latter model versions and with some consideration of how recent model developments may impact results presented here.

- *For example, no reference is made to the inclusion and evaluation of primary marine organic aerosol in later model versions. Additionally, sea salt density has been updated, as has deposition velocities via land surface representations, both of which would affect the sensitivity test results.*

Thank-you for highlighting these changes in the model. We note that on Line 689 we did make reference to the fact that 'Switching on PMO and re-scaling DMS brings ACCESS inline with more recent versions of the UM global atmosphere configurations' however we take the point that this should have been made more clear. We have now done this in the methods section as well.

Line 158: 'The DMS flux is scaled by a factor of 1.7 to take into account the lack of PMO, which are not switched on by default (Mulcahy et al. 2020). This is different in later versions of GLOMAP-mode, which returns DMS to a scaling of 1 and turns of PMO (Mulcahy et al. 2020).'

As for the changes to the sea salt density, we note that this change was adopted in GLOMAP-mode in ACCESS-CM2 (and therefore ACCESS-AM2), though this has not been documented. We will rectify this in the methods section.

Line 160: 'SSA emission fluxes are calculated using the wind-speed parameterization source function developed by Gong (2003), and include the updated sea salt density as per Mulcahy et al. (2020).'

We have not found evidence in the model branches/code that the updates to the deposition velocities were included in this version of ACCESS, so we have made comment as the Reviewer suggests about how this may impact our results in our Discussion section.

Line 656: 'A later update to the GLOMAP-mode dry deposition velocities has lead to increased coarse-mode deposition velocities that reportedly impact the sea-salt aerosol distribution (Mulcahy et al. 2020). We speculate that this update may reduce the magnitude of impact of our SSA changes on the CCN.'

Figures:

- *Font size in figures is sometimes too small. Additionally, thicker lines with better color contrast, or some other method, is needed to distinguish between simulations.*

We have increased the line thickness and font size in the figures. We have left the colours as is in order to ensure they are colour blind safe, and note that this is why we have also used different line styles.

- *On line 511, the authors state they have evaluted other cloud properties, which is essential to make a complete analysis of the impact of the sensitivity tests on aerosol, clouds and aerosol-cloud interactions. Equivalent figures should be included in a supplement, so the reader can interpret the wider effects themselves.*

We take the point from this Reviewer, and note that this is something we discussed extensively amongst the authors. We also note the opposing view from Reviewer 2, to remove the LWP discussion as it is not evaluated against observations. In this light, we have decided to leave the text as it is for two reasons: a) the cloud-radiative biases in ACCESS-AM2 have been well documented in our previous work, and b) we do not want to take away from the main point of this study, which is the aerosol evaluation. We acknowledge that we have made some strong statements (which this Reviewer disagrees with) about the climatic impacts of the changes to the aerosol scheme. We have lessened these statements to better reflect what our results are, and to highlight them more clearly.

References:

Some additional references that have been overlooked include:

- *Schutgens et al. (2017) doi.org/10.5194/acp-17-9761-2017 in section 3.4.1*
-

We thank the Reviewer for bringing this work to our attention, we have now included it in the methods section:

Line 364: 'Schutgens et al. (2017) recommend collocating model and observational data at hourly intervals to reduce representation error. However, for these simulations this approach was not feasible. '

- *Additional literature evaluating more sophisticated NPF schemes and the climatic effects of those schemes.*

We note that more sophisticated NPF schemes are briefly mentioned in the Introduction. We have now also included the following in our Discussion section:

Line 615: 'Most global atmosphere models use classical nucleation theory involving binary NPF, however, more complex ternary or ion nucleation parameterisations have also been developed. For the Antarctic, ion nucleation of sulfuric acid with ammonia (sourced from sea bird colonies) has been suggested to be an important pathway for nucleation (Lee et al. 2019), implying that a more complex NPF scheme could benefit this region. However, significant updates to the chemistry in ACCESS would be required to include such sources.'

- *The final sentence in section 4.2*

This sentence has been removed.

- *Model structural changes implemented after this model version, particularly where they may affect interpretation of results here (e.g. https://gmd.copernicus.org/articles/13/6383/2020/)*

We have addressed the structural changes as per the previous comment in Section 3.1.1.
Spelling and syntax:

- *Line 101: '(GLOMAP)' over-used*

  We have rephrased this sentence as:

  Line 113: '... which includes the Global Model of Aerosol Processes (GLOMAP)-mode aerosol scheme... '

- *Line 149: 'volcinic-sourced'*
  Hyphen added
- *Equation 7: Numerator should be 'CHL'*
  Numerator fixed
- *Line 261: Remove '/,'*
  LaTeX syntax fixed
- *Line 290: SYO not defined*
  We have removed all reference to SYO throughout text
- *Line 310: inline*
  All instances of 'inline' have been change to 'in line'
- *Section 3.4.2: This could easily be a single paragraph*
  Has been made one paragraph
- *Line 353: missing 'is'*
  Fixed
- *Line 605: bis*

Fixed