Response to Reviewer #1

Reviewer comments are in **bold** and the authors' responses are in blue.

We thank the reviewer for their thorough and constructive evaluation of our manuscript. In response, we have removed unsupported assertions about nitrate and dust biases canceling between scenarios, clarified that nitrate–ammonium interactions are not represented in CESM2(WACCM6), and explicitly noted that WACCM's dust overestimation may inflate the prominence of dust in Fig. 2. We have also added the previously missing caveat in Sec. 2.1 and revised the title to avoid ambiguity around the term "deposition."

First, the authors argue that the lack of nitrate aerosol is unlikely to cause a systematic bias because they are comparing ARISE-SAI-1.5 and SSP2-4.5 (lines 111-113). I do not understand this argument. Surface-level sulfate aerosol competes with nitrate aerosol, and can indeed displace it; furthermore nitrate aerosol/gas partitioning will change in response to near-surface temperature whereas near-surface sulfate partitioning is insensitive to temperature. This feels particularly significant given that SAI is expected to change, among other things, near-surface temperature. The lack of nitrate or ammonium aerosols in WACCM means that these interactions are not captured, and it is very difficult to say what their implications would be; consider, for example, work such as Tai et al. (2012) which suggests that nitrate specifically exhibits a different trend with climate change than other constituents. The statement that nitrate is "standard in some regional air quality models" (line 105), implying that the absence of nitrate in CESM2-WACCM for an air quality assessment is normal, also risks being somewhat misleading considering that CESM2-WACCM was the only model out of 6 global Earth system models in a recent intercomparison which lacked an explicit representation of nitrate gas/aerosol partitioning (He et al., 2025). I would recommend removing assertions regarding the size of the effect entirely and simply stating that WACCM does not include this interaction, as there is no clear evidence that the associated errors are negligible.

We thank the reviewer for this important clarification. We agree that nitrate–ammonium thermodynamics and nitrate–sulfate competition can significantly affect PM2.5 levels and that the implications of omitting nitrate aerosol cannot be assumed negligible. In response, we have removed assertions that the lack of nitrate aerosol is unlikely to cause a systematic bias and that nitrate is "standard in some regional air quality models".

Similarly, the fact that the authors are comparing these scenarios does not seem to support the argument regarding WACCM's known dust bias that "[b]ecause the same definition of PM2.5 is applied consistently across both scenarios, any systematic bias in the representation of dust is expected to cancel out when examining the relative effects of SAI". The problem with this argument is that, if dust exposure is overestimated by (say) 50%, then so too is the absolute change in dust exposure from which mortality estimates are calculated. This calls into question conclusions about dust being a primary driver (e.g. line 353), as well as figures such as Figure 2 which show dust as being the dominant contributor to changes in PM2.5 in many locations. I would recommend at least considering how the conclusions would be affected if the change in dust is

overestimated relative to changes in other constituents, perhaps using work such as Hancock's to determine the likely overestimate in dust concentrations (and therefore in estimates of absolute changes in dust between scenarios). If anything this supports the authors' conclusions that changes in PM2.5 due to SAI are expected to be small.

We agree that the previous wording may have overstated the extent to which dust-related biases might cancel when comparing ARISE-SAI-1.5 and SSP2-4.5. As the reviewer notes, if dust concentrations are systematically overestimated, then the absolute magnitude of changes in dust (and thus dust-attributable PM2.5 changes and mortality) may also be overestimated. This uncertainty has implications for interpreting Fig. 2 and for statements identifying dust as a primary driver of PM2.5 differences. In response, we have revised the manuscript to explicitly acknowledge, in the discussion accompanying Fig. 2, that WACCM's dust overestimation could inflate both absolute dust burdens and their apparent contribution to PM2.5 changes. In particular, we clarify that dust's prominence in Fig. 2 should therefore be interpreted with caution, and that dust may appear as a dominant PM2.5 constituent in part due to this known model bias. We also emphasize (as the reviewer points out) that this uncertainty reinforces, rather than contradicts, our broader conclusion: namely, that the air-quality impacts of SAI are small compared with projected policy-driven improvements, and that the absolute contribution of sulfate is modest in the context of future air-quality change.

On a more minor note, the authors state with regards to the dust bias that they "have included text in the manuscript to highlight this caveat". I cannot find any such additional text in the manuscript, so I would recommend that the authors make clear this caveat.

We thank the reviewer for pointing this out. We made sure this time the caveat is included in the discussion of the model limitations under Sec. 2.1.

Finally, I realised that the title might be confusing to some readers. As it stands, it reads "Air quality impacts of stratospheric aerosol injections are likely small and mainly driven by changes in climate, not deposition". However lines 413-416 state that "regional changes in PM2.5 concentrations and the corresponding health impacts are mainly driven by shifts in precipitation patterns and/or circulation, which affect the wet removal of non-sulfate species such as dust and secondary organic aerosols". I fear that the term "deposition" in the title is liable to lead to confusion; I assume the authors mean "settling of injected stratospheric aerosol to the surface", but an air quality expert may instead read it as "wet deposition (i.e. precipitation-related scavenging)". I would recommend modifying the title to make clear exactly what form of deposition is intended, e.g. "...changes in climate, not descent of stratospheric aerosol to the surface" (or ideally something less wordy).

We have changed the title of this manuscript to “Air quality impacts of stratospheric aerosol injections are likely small and mainly driven by changes in climate, not aerosol settling” to address the reviewer’s comments.

Response to Reviewer #2

Reviewer comments are in **bold** and the authors' responses are in blue.

We thank the reviewer for their thorough and constructive feedback, which has improved the clarity of the manuscript. In response, we clarified ambiguous language, removed redundant definitions, standardized terminology, and strengthened our discussion of ensemble robustness.

Introduction: Line 28: "release of precursors such as sulfur dioxide" – precursors to what? Admittedly, readers are not likely to be confused by this, but the sentence should be made precise.

We thank the reviewer for pointing out this ambiguity. We have revised the sentence to clarify that sulfur dioxide (SO2) is the precursor to sulfate aerosols.

Model description: Line 93: CESM2(WACCM6) is already defined (and used) as an abbreviation earlier in the text (lines 39-40, albeit without version).

We have removed the repeated definition in the Model Description section and now simply refer to CESM2(WACCM6) using the abbreviation established earlier in the Introduction.

Line 109: Suggest moving "(Wen et al, 2023; Wei and Tahrin, 2024)" to end of sentence in line 110, considering this second part contains the findings of said papers.

The citations have been moved to the end of the sentence.

Line 131: "with the aim of maintain[ing]"

The text has been revised accordingly.

Results: Prior to "In Fig. 1, …" I would highly recommend the authors to add a topic sentence, with the main conclusion of this paragraph. This is completely optional, but paragraph 3.1 currently reads like a summation of figures first and foremost, and less like an active interpretation of these findings. A topic sentence with the main findings before diving into the figures would be a welcome addition here.

To improve the clarity and flow of Section 3, we have moved the original first paragraph of Section 3.1 out of the subsection and placed it at the beginning of the Results section. This paragraph provides an overarching explanation of the three-way comparison used throughout the analysis and therefore serves more appropriately as general guidance for how the results should be interpreted, rather than as an introductory sentence to Section 3.1.

Because this paragraph establishes the interpretive framework for the analysis, we do not believe it would be appropriate to summarize the findings of Section 3.1 & 3.2 before presenting the underlying diagnostics, mechanisms, and figures that support those findings.

Response to Reviewer #2

The results section often mentions ensemble spread but does not clearly identify which results are robust across the ensemble and which are dominated by variability. For example, the hemispheric ozone asymmetry appears as a strong and consistent feature, whereas several PM2.5 patterns do not. Clarifying this distinction will help readers assess confidence in each result.

Related to this, many maps contain extensive stippled regions indicating non-significant differences. In some cases, the text still describes spatial patterns along these regions. It would be helpful to explicitly state when results are not statistically significant rather than implying interpretability from noisy patterns.

In the revised manuscript, we now explicitly distinguish between ensemble-robust features and those that exhibit substantial internal variability. Specifically:

- We emphasize that the hemispheric ozone asymmetry is a strong, statistically significant, and consistent feature across all ensemble members.
- For PM2.5, we now state clearly that many spatial patterns show large ensemble spread and are therefore less robust, with only a few regions exhibiting statistically significant or ensemble-consistent changes.
- We also updated the description of Fig. 2 to note how stippled areas reflect limited ensemble agreement, but do not affect the overall conclusions that can be made that non-sulfate species dominate PM2.5.
- We have revised the text to explicitly note that many of the PM2.5-related mortality changes occur in regions where internal variability dominates, consistent with the broader PM2.5 spatial pattern being statistically insignificant across much of the globe.

Line 216: "ITCZ" is defined, but not used again, so this (ITCZ) can be omitted.

"ITCZ" has been removed since it is not used again.

Figure 3: I suggest being consistent and writing out "percent" all the time, or even use "relative" instead of switching between "percent" and "%".

We have revised the figure caption and all associated text to use consistent terminology. Specifically, we now write out "%" throughout the manuscript for clarity and uniformity.