This manuscript addresses a worthwhile and relatively understudied aspect of MCB: the consequences of MCB-induced climate changes for surface ozone. The authors employ a three-stage experimental framework to first determine the sea salt aerosol injection rate required to achieve a target radiative forcing, simulate the coupled climate response, and subsequently examine the resulting changes in surface ozone using prescribed SST and SIC fields in atmospheric chemistry simulations. The study further applies machine-learning-based bias correction to the simulated ozone concentrations and uses SHAP analysis to identify meteorological variables associated with the regional ozone responses.

However, I have several concerns regarding the robustness of the reported ozone responses, the interpretation of the machine-learning results, and the extent to which the proposed physical mechanisms are supported by the analysis. In particular, the bias correction substantially modifies the magnitude of the simulated ozone responses, yet its reliability under the future MCB-modified climate is not adequately established. The subsequent statistical attribution of meteorological drivers also raises questions about the independence of the relationships identified. More generally, while the manuscript documents regional changes in ozone concentrations, I find that some of the conclusions regarding their physical drivers and broader air-quality implications extend beyond what is demonstrated.

I believe the manuscript could make a useful contribution to the literature, but these issues should be addressed before it can be considered for publication. My specific comments are provided below.

Major comments

1) The authors train LightGBM models using historical observations and CESM1 simulations from 2014–2019 and subsequently apply these models to the future SSP5-8.5 and MCB simulations. However, it is unclear whether the statistical relationships learned from historical conditions remain valid under the future climate, particularly when MCB introduces changes in temperature, humidity, atmospheric circulation, and chemical conditions. Furthermore, the magnitude of the bias correction is substantially larger than the reported MCB-induced ozone responses. For example, the authors report regional mean corrections of approximately 20–40 μg m-3 (Lines 312–321), whereas the resulting changes in ozone concentrations are generally less than 1 μg m-3. Although large corrections to the baseline ozone concentrations do not necessarily imply large errors in the simulated ozone response, the changes in the response after correction warrant closer examination. The European mean response changes from +0.33 μg m-3 before correction to +0.01 μg m-3 afterward. The response over eastern China also changes substantially, from -1.14 to -0.64 μg m-3. This raises the question of how sensitive the conclusions are to the particular bias-correction model employed.

   The historical simulations used for training include nudging of the horizontal winds toward MERRA-2 reanalysis (Lines 165–167), whereas the future chemistry experiments

are not described as being similarly constrained. This potential difference in meteorological conditions between the training and application datasets should be discussed.

The authors acknowledge some of these limitations in Lines 403–420, but given the central role of the corrected results in the manuscript, I believe a more quantitative assessment is warranted. In particular, evaluating the sensitivity of the corrected ozone responses to the training procedure, consider temporally or spatially independent validation, and examine whether the future predictor values fall outside the historical training distributions.

Most importantly, improved agreement with historical ozone observations does not necessarily indicate improved accuracy in the simulated ozone *response* to MCB. The authors should provide stronger justification for interpreting the corrected changes as more reliable than the uncorrected model results.

2) The authors use a second set of LightGBM models to identify the meteorological variables associated with the spatial variability of the bias-corrected ozone response. While this is potentially useful, I have concerns about the interpretation of the resulting SHAP analysis.

   The bias-correction models use meteorological variables, among other predictors, to estimate corrections to the simulated ozone concentrations. The second set of LightGBM models then uses changes in overlapping meteorological variables to explain the resulting bias-corrected ozone response.

   Consequently, some of the relationships identified in the SHAP analysis may reflect the dependence introduced by the bias-correction procedure itself, rather than independently demonstrating the physical processes responsible for the ozone changes. This issue is particularly important because the manuscript uses these results to support conclusions regarding the relative importance of moisture, atmospheric transport, and other meteorological processes.

   I would suggest repeating the attribution analysis using the uncorrected CESM1 ozone response and comparing the results with those obtained using the bias-corrected response. This would help establish whether the identified relationships are robust to the bias-correction procedure.

   Additionally, the five predictor groups contain different numbers of variables, some of which are strongly correlated. For example, the moisture group includes specific humidity, relative humidity, and precipitation, whereas the thermal group includes only surface temperature. Summing the absolute SHAP values within these groups does not necessarily provide a direct comparison of their independent physical importance.

   The authors appropriately acknowledge that SHAP identifies statistical associations rather than causality (L 356–358). Nevertheless, the interpretation of Figure 7 should be

more cautious, particularly when discussing the relative importance of the different physical processes.

3) The manuscript attributes the spatially heterogeneous ozone responses to changes in temperature, moisture, cloud cover, radiation, boundary-layer mixing, and atmospheric transport. While these explanations are physically plausible, several of the proposed mechanisms are not sufficiently demonstrated by the analysis presented. For instance, the authors attribute positive ozone anomalies over tropical and subtropical regions to atmospheric drying and reduced water vapor-dependent chemical loss (L 263–269). Over eastern China, ozone reductions are attributed primarily to cooling and suppressed photochemical production, whereas over South Asia, they are linked to increased cloud cover and reduced radiation.

   The chemical production and loss diagnostics presented in Figure 5 provide useful supporting information. However, ozone concentrations are determined by the combined effects of chemical production, chemical loss, atmospheric transport, mixing, and deposition. Changes in chemical production or loss alone do not necessarily establish their relative contributions to the resulting concentration changes.

   The authors should strengthen the mechanistic interpretation by examining changes in net chemical ozone tendencies, rather than primarily relying on the spatial correspondence between ozone and meteorological anomalies. Where possible, additional information on transport or deposition would also help determine whether the proposed chemical mechanisms adequately explain the simulated changes.

   A comprehensive ozone budget analysis is not necessarily required, but a more quantitative examination of the dominant processes in the regions emphasized by the manuscript would strengthen its scientific contribution. Otherwise, the discussion should more clearly distinguish between mechanisms supported by the model diagnostics and those that remain plausible explanations.

4) Many of the regional mean ozone responses reported in this study are relatively small. Following bias correction, the summer mean changes are -0.33 μg m-3 over China, +0.01 μg m-3 over Europe, and +0.12 μg m-3 over the United States. Given the small magnitude of these responses, the statistical robustness of the results deserves greater attention.

   The figure captions indicate that t-tests are used to evaluate the significance of spatial changes, and Figure 6 presents 95% confidence intervals for regional means. However, the methodology used to calculate these confidence intervals and the treatment of temporal variability are not sufficiently explained. In particular, it would be helpful to clarify how the independent samples are defined, whether temporal autocorrelation is taken into account, and how the variability among the three coupled ensemble members is represented in the prescribed SST and SIC forcing used for the atmospheric chemistry

experiments. It is also unclear whether the confidence intervals for the corrected ozone responses incorporate uncertainty associated with the bias-correction procedure.

The authors should provide a more detailed description of their statistical methodology and distinguish between changes that are statistically robust and those that are small relative to the associated uncertainty. This is especially important for the regional conclusions, where some of the reported changes are close to zero.

5) The three-stage experimental design is useful for separating the influence of MCB-induced climate changes from the direct chemical effects of sea salt aerosol injection. However, I think the authors should more carefully define what is meant by the "climate-mediated" ozone response.

   In Stage III, the authors prescribe SST and SIC fields from the coupled simulations while omitting the additional sea salt aerosol injection. This approach captures ozone responses associated with the changes in SST and SIC and the resulting atmospheric adjustments. However, it does not necessarily capture all indirect physical effects of MCB. For example, the atmospheric chemistry experiments do not explicitly reproduce the aerosol-induced changes in cloud properties and radiation that occur during MCB deployment. Rapid atmospheric adjustments associated with these perturbations may also influence ozone chemistry independently of the SST and SIC response.

   The study more precisely isolates the SST- and SIC-mediated component of the ozone response, rather than the complete indirect response to MCB. This distinction should be clarified throughout the manuscript, particularly in the Abstract, Introduction, and Conclusions. This would better communicate the scope of the experimental framework and avoid implying that all climate-mediated effects of MCB have been accounted for.

6) The authors calibrate the sea salt injection rate to achieve an effective radiative forcing perturbation of -1.0 W m-2, corresponding to the stated difference between SSP5-8.5 and SSP2-4.5 in 2050. However, this calibration is performed using prescribed present-day SST and SIC conditions, whereas the coupled MCB simulations are conducted under year-2050 SSP5-8.5 forcing.

   Since aerosol–cloud interactions and the resulting radiative forcing may depend on the background climate state, it is unclear whether the calibrated forcing perturbation remains equivalent under the future conditions used in the coupled simulations. I suggest that the authors discuss this uncertainty and, if possible, provide an estimate of the realized forcing under the future climate conditions. Additionally, the coupled simulations are integrated for 300 years with year-2050 forcing held constant. The resulting climate response may reflect the long-term adjustment under fixed 2050 conditions rather than the transient response expected in 2050.

7) The authors should expand their discussion of previous studies examining the effects of solar geoengineering on tropospheric and surface ozone. In particular, Xia et al. (2017)

and Wang et al. (2026) investigated ozone responses under stratospheric aerosol injection (SAI), including the roles of climate-mediated chemical changes and atmospheric transport. The authors also cite Bednarz et al. (2025), which examined stratospheric and tropospheric ozone responses to MCB. Given the close relevance of this study, a more detailed comparison of the findings, particularly regarding the mechanisms governing tropospheric ozone changes, would help clarify the additional insights provided by the present work. More broadly, comparing these studies would help establish how the magnitude, spatial patterns, and underlying mechanisms of ozone responses under MCB may differ from those under SAI.

Furthermore, the authors may find it useful to discuss their findings in the context of the recently developed G6-1.5K-MCB GeoMIP protocol (Hirasawa et al., 2026). This study highlights the substantial intermodel differences in regional climate responses to MCB. Given that the ozone changes identified in the present manuscript are largely associated with changes in meteorological conditions, it would be worthwhile to discuss how sensitive the reported ozone responses might be to the choice of climate model and MCB deployment strategy. This would also help clarify the extent to which the findings can be generalized beyond the specific experimental configuration considered here. Liu et al. (2022) previously investigated machine-learning-based bias correction of future surface ozone projections. Given the importance of bias correction to the results presented here, this study should also be acknowledged, particularly when discussing how bias correction affects the magnitude and spatial patterns of projected ozone changes.

Minor comments:

1) L 19-21: Are the injection rates of 11 Tg yr-1 applied to each of the six regions, corresponding to a total injection of 68.4 Tg yr-1 (as indicated in the abstract)?
2) L 23–24 and 227–230: How is the reported 9.7% increase in Arctic sea ice concentration calculated, and does it represent a relative percentage change or an absolute difference in percentage points?
3) L 30–32: I find the concluding statement of the abstract somewhat overstated relative to the results presented. The study primarily provides a quantitative description of regional changes in summertime surface ozone concentrations, many of which are relatively small. However, the authors do not establish whether these changes are sufficiently large or robust to have meaningful implications for regional air quality. Furthermore, the analysis considers only the climate-mediated ozone response under a single MCB scenario and does not account for other potentially important chemical effects. This statement should be modified to better reflect the scope of the analysis and the strength of the conclusions that can be drawn from the results.
4) Since “sea salt aerosols” is presented in L 49 prior to L 66 when the acronym SSA is introduced, you can introduce the acronym SSA earlier in L 49.
5) L 138-145: Future biogenic VOC emissions are prescribed using offline inputs and soil NOx emissions are kept at present-day levels. However, both biogenic VOCs and soil NOx emissions may respond to MCB-induced changes in temperature, radiation, and

soil moisture. Therefore, holding these emissions fixed excludes potentially important climate-mediated pathways influencing ozone production. This limitation and its potential implications should be acknowledged and discussed.

6) Section 2.4 and Table 1: Please clarify the integration length, spin-up period, and analysis period for the atmospheric chemistry simulations in the main text. It would also be useful to specify more clearly how the SST and SIC fields from the coupled ensemble members are used to construct the prescribed boundary conditions.
7) L 153-156: Include a citation to support this sweeping statement.
8) L 227: Does the reported localized warming of approximately 1.29°C refer specifically to the SST increase around Greenland? As currently phrased, the statement suggests that localized warming is widespread, whereas Figure 2 shows warming primarily around Greenland, with cooling across most other regions. I suggest clarifying the geographical extent of this warming and focusing the discussion on the broader spatial pattern of SST responses, rather than making a general statement about localized warming.
9) L 228: suggest reporting the global mean change in sea ice concentration (SIC), along with the overall range of SIC changes. This would provide a more informative summary of the sea ice response than the statement that local increases exceed 16.5%, particularly since these localized changes are difficult to discern from Figure 2. Reporting the global mean change and the range of regional changes would help readers better understand the magnitude and spatial variability of the SIC response to MCB.
10) L 229-230: The statement attributing the Northern Hemisphere SIC response to differences in ocean heat capacity, land-sea configuration, and ice-albedo feedback seems overly broad and is not sufficiently supported by the analysis presented. While these factors may contribute to hemispheric differences in sea ice responses, their relative roles have not been investigated in this study. Either provide additional evidence to support this interpretation or temper the statement to avoid attributing the simulated SIC changes to mechanisms that have not been explicitly examined.
11) L 264–266**:** The statement that the meteorological responses are characterized by widespread surface cooling and regionally varying changes in moisture, cloud cover, boundary-layer structure, and atmospheric circulation is rather general. While this description is consistent with Figure 4, it does not provide much insight into the specific spatial patterns of these responses.
12) L 327–329 and 335-337: The preceding discussion does not necessarily demonstrate the importance of incorporating regional air quality responses into assessments of MCB deployment, but rather highlights the sensitivity of the simulated ozone responses to the bias-correction procedure. In particular, the magnitude and spatial patterns of the ozone changes differ between the raw and bias-corrected simulations, suggesting that the methodology used to correct mode biases can substantially influence the interpretation of the results. While the subregional differences also demonstrate that regional averaging can obscure spatial heterogeneity, this does not necessarily establish meaningful implications for air quality, particularly given the relatively small magnitude of the reported ozone changes. These statements should be tempered and instead focus on what the results directly demonstrate: the importance of bias correction and the

potential limitations of regional mean estimates in characterizing spatially heterogeneous ozone responses.

13) L 340–360 and Figure 7: I suggest consistently using terms such as "predictive importance" or "statistical association" rather than "physical contribution" when interpreting the SHAP analysis. In addition, the foundational SHAP methodology should be appropriately referenced, for example, Lundberg and Lee (2017).
14) Section 3.3: It would be helpful to report the regional ozone responses as percentages of the corresponding baseline concentrations, in addition to the absolute changes. This would provide further context for assessing the relative magnitude of the changes across regions.
15) Section 3.3 and Conclusions: The authors frequently discuss the implications of the ozone changes for regional air quality. However, the analysis focuses on summertime mean ozone concentrations. Air-quality impacts are often evaluated using metrics such as the maximum daily 8-hour average ozone concentration, which may respond differently from seasonal mean concentrations. While I do not think an additional analysis of ozone extremes is strictly necessary, the authors should acknowledge this distinction when discussing the broader air-quality implications.
16) Section 4: The authors acknowledge several important limitations of the analysis, including the uncertainty associated with bias correction and the statistical nature of the SHAP attribution. Some of these limitations directly affect the interpretation of the main results and should be addressed more explicitly in the relevant Results sections, rather than being discussed primarily as limitations for future work.
17) It would be helpful to add another row of figures in Figure S2 showing the seasonal O3 differences between CTRL and MCB.
18) Figure 6: How are the 95% confidence intervals in the error bars calculated? Do these intervals represent temporal variability, uncertainty in the mean response, or another source of uncertainty?

References:

Bednarz, E. M., Haywood, J. M., Visioni, D., Butler, A. H., and Jones, A.: How marine cloud brightening could also affect stratospheric ozone, Sci. Adv., 11, eadu4038, https://doi.org/10.1126/sciadv.adu4038, 2025.

Hirasawa, H., Henry, M., Rasch, P. J., Wood, R., Doherty, S. J., Haywood, J., Wong, A., Lamarque, J.-F., Brody, E., and Wang, H.: G6-1.5K-MCB: Marine Cloud Brightening scenario design for the Geoengineering Model Intercomparison Project (GeoMIP) in CESM2.1, E3SMv2.0, and UKESM1.1, Geosci. Model Dev., 19, 3257–3283, https://doi.org/10.5194/gmd-19-3257-2026, 2026.

Liu, Z., Doherty, R. M., Wild, O., O'Connor, F. M., and Turnock, S. T.: Correcting ozone biases in a global chemistry–climate model: implications for future ozone, Atmos. Chem. Phys., 22, 12543–12557, https://doi.org/10.5194/acp-22-12543-2022, 2022.

Lundberg, S. M. and Lee, S.-I.: A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst., 30, 4765–4774, https://papers.nips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html, 2017.

Wang, C., Visioni, D., Chua, G., and Bednarz, E. M.: Air quality impacts of stratospheric aerosol injections are likely small and mainly driven by changes in climate, not aerosol settling, Atmos. Chem. Phys., 26, 1339–1357, https://doi.org/10.5194/acp-26-1339-2026, 2026.

Xia, L., Nowack, P. J., Tilmes, S., and Robock, A.: Impacts of stratospheric sulfate geoengineering on tropospheric ozone, Atmos. Chem. Phys., 17, 11913–11928, https://doi.org/10.5194/acp-17-11913-2017, 2017.