Responses to referee 1

General comments:

In this study, the authors explore the effect of the Holuhraun eruption on cloud properties on a kilometer-scale simulation within the UK Met Office Unified Model. In their setup, the authors employ a detailed cloud microphysical scheme and an interactive aerosol module, making their simulations well-suited to study aerosol-cloud interactions resulting from the Holuhraun eruption. They furthermore nicely outline the importance of considering meteorology and background aerosol conditions in regions that are not directly affected by volcanic aerosol when determining the effect of volcanic aerosols on clouds.

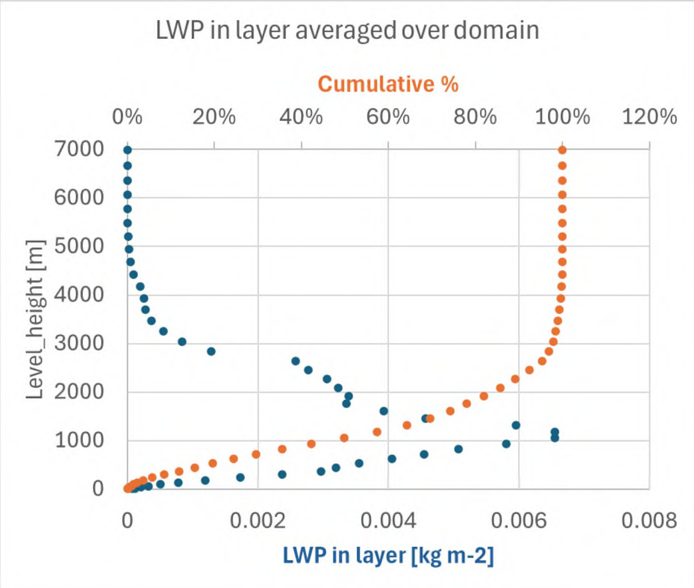
The manuscript is logically structured and well written and merits publication provided that the following comments are addressed.

Specific comments:

- P6, L210-211: Could the authors clarify the rationale behind using a 5 km cloud top threshold? Why was this specific threshold chosen? Given that cloud phase is strongly temperature-dependent, a threshold more directly related to cloud thermodynamics, such as cloud top temperature, may be more appropriate.

We thank the reviewer for this comment. The cloud top threshold of 5 km was chosen as a practical criterion to separate lower tropospheric liquid clouds from upper tropospheric ice clouds such as cirrus. A domain mean height profile of liquid water content shows that more than 99% of the liquid water resides below 4 km (see the plot below). Therefore, a 5 km threshold effectively captures the relevant population of marine liquid clouds, and we do not expect the precise choice of this upper limit to strongly affect our results. Furthermore, the same threshold was applied consistently to both satellite observations and model simulations, so any potential bias introduced by this choice is mitigated in the comparisons. We added the following sentence in the main text;

This 5 km threshold was chosen to separate lower-tropospheric liquid clouds from upper-tropospheric ice clouds such as cirrus. In our simulations, more than 99% of liquid water resides below 4 km, so varying this threshold within a reasonable range (e.g. 4–6 km) would not substantially affect the results.

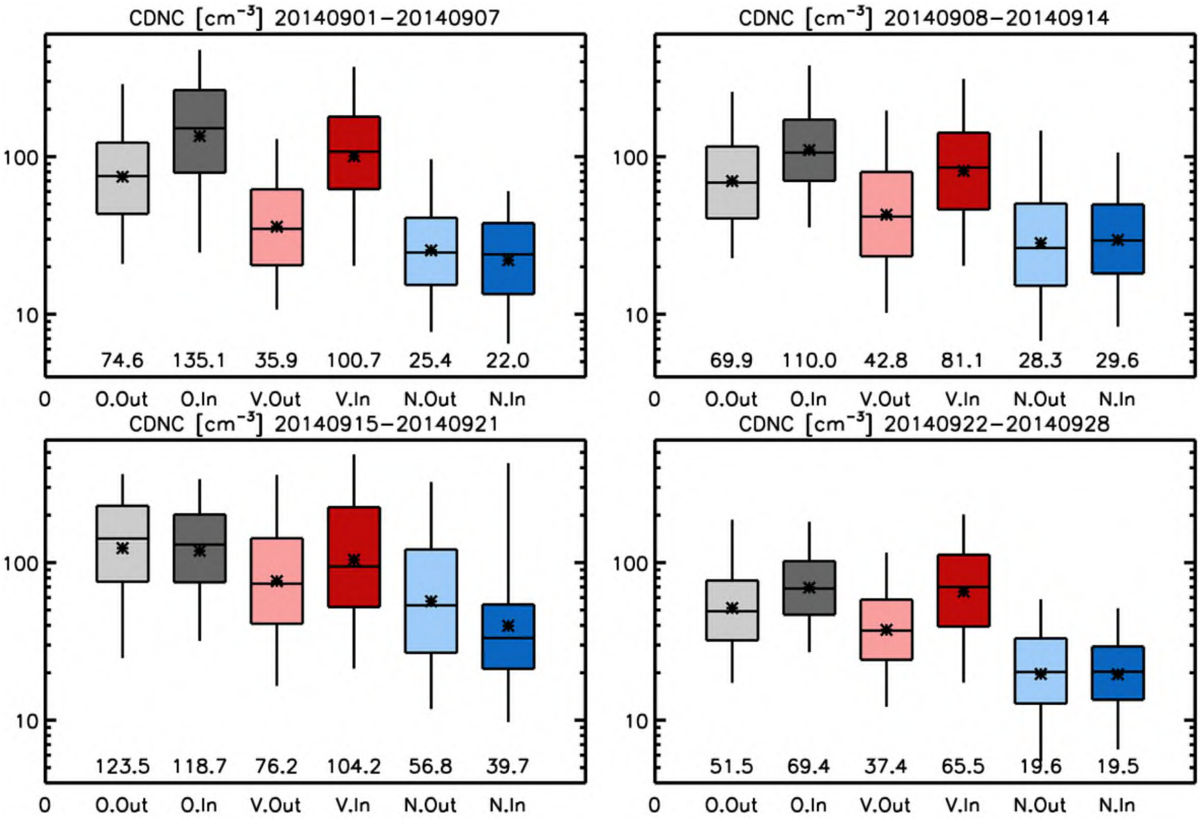


• P7, L226-229: The bounding boxes used for the satellite and model data differ in size and shape, which likely introduces biases when comparing in-plume and out-of plume cloud properties. Since meteorological conditions and, therefore, cloud characteristics vary spatially, using different out-of-plume areas may influence the interpretation of perturbation signals. Furthermore, the authors provide little information on how they define the rectangular boundary that envelops the in-plume regions (I assume it is similar to that in Peace et al., 2024). Could the differences reported in Table 4 stem from different areas of in- and out-of-plume? This is for sure the case for the data from Haghighatnasab et al. (2022), and looking at Peace et al. (2024), the regions are also different.

The bounding boxes are defined using satellite-retrieved and model-simulated column SO_2 burdens, which differ from each other. These reflect the differences in atmospheric transport and dispersion between the real atmosphere and the model. Therefore, the boxes are internally consistent within each dataset and are suitable for comparing the effects of volcanic plume in a self-consistent manner. Using identical geographic regions for model and observations would lead to include plume effects in the out-of-plume regions and vice versa.
We have expanded the description of choice of regions as follows:
“In-plume” and “out-of-plume” regions were determined by selecting the minimum and maximum east-west and north-south extents of the plume, following the methodology of Peace et al. (2024).

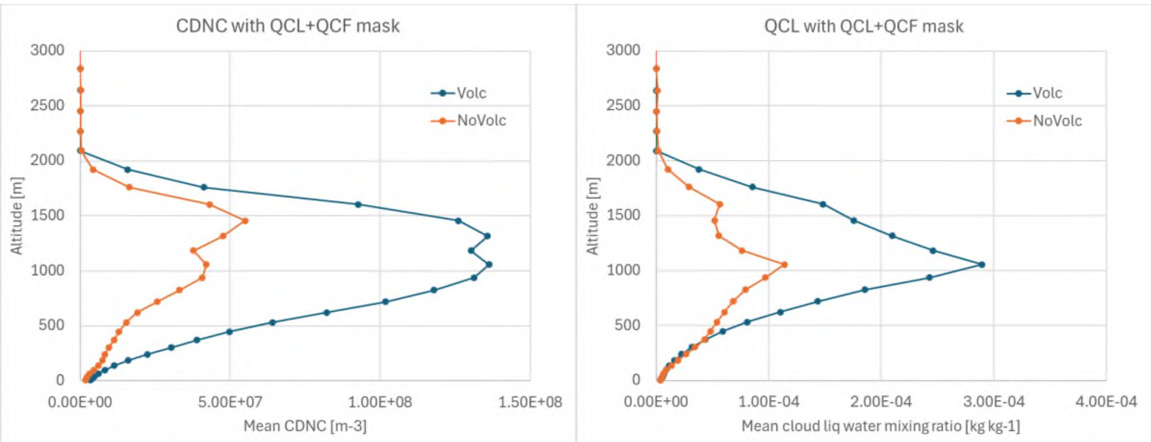
• P7, L231-232: The threshold of 0.5 DU for defining in-plume regions in the model was chosen to align spatial patterns with satellite data. However, this lower threshold could lead to an underestimation of sulfate burden and, consequently, cloud condensation nuclei (CCN) and cloud droplet number concentration (CDNC) in the model. It would be informative if the authors could show the sensitivity of CDNC to different threshold choices, for example, what changes occur if a 1 DU threshold is applied? Could this partially explain the remaining CDNC underestimation, even after retuning?

We share the referee's concern about the difference in thresholds applied to satellite observations and model simulations, which could affect the analyses. To address this, we generated plots equivalent to Fig. 4 using a 1 DU threshold for the simulation data instead of 0.5 DU.
These plots show slightly higher CDNC inside the plume, but we confirm that this difference does not significantly affect our analyses or conclusions.
We added these plots in Fig. S2 and added this sentence in the main text. "While using a 0.5 DU threshold provides better spatial agreement between the satellite and model aerosol fields, using a 1.0 DU threshold does not qualitatively or quantitatively (up to 18% increase in Volc.In CDNC) change the results (see figure S2)."



• P7, L246-248: The authors claim that their model-derived CDNC is comparable to MODIS-derived CDNC, yet no supporting analysis is provided. MODIS CDNC retrievals are predominantly sensitive to cloud top, while for the model, a liquid-water weighted vertical average is computed. Although this weighting gives some emphasis to upper layers, it likely results in lower CDNC values compared to satellite retrievals. I know that it is intricate to perform a fully definition-aware comparison between model and satellite observations (e.g., using a satellite simulator like COSP). I would, nevertheless, ask the authors to somehow show that their CDNC values are comparable to the satellite-derived ones. Alternatively, would it be feasible to extract model CDNC at cloud top for a more definition-consistent comparison?

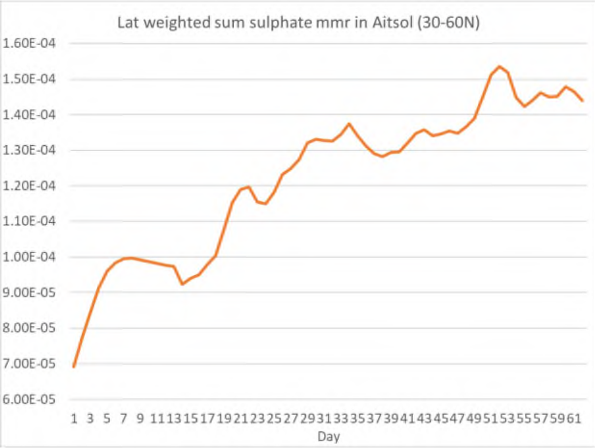
To demonstrate that using liquid water content weighted CDNC to compare to MODIS is acceptable we show some analysis where we have selected columns where condensate in any grid box below 2 km greater than 0.5 g m-3 with clear sky above. This picks out columns that are predominantly stratus cloud and eliminates the deeper frontal clouds that will present ice to MODIS. Composite plots of CDNC and LWC in the volcano plume area with volcano on and off show that the LWC weighting will emphasise where the peak grid mean CDNC values are for comparison with MODIS.
The MODIS retrieval of CDNC assumes a constant CDNC throughout the depth of the cloud, rather than retrieving cloud top CDNC per se. However, the CDNC retrieval is very sensitive to the effective radius retrieval, which is weighted towards cloud top (in terms of optical depth from cloud top). It also assumes a linear increase in liquid water content (LWC) with height, so that LWC is highest at cloud top under this assumption.
Our approach of weighting the model CDNC by LWC is a compromise to ensure that the vertical levels with high LWC, which are most important for the cloud optical depth and radiative effect, contribute most to the mean CDNC value. It also reduces the influence of spurious thin (low LWC) cloud levels.
If the model clouds behave similarly to the assumptions made for the MODIS CDNC retrieval, then the exact vertical weighting would not matter much, since CDNC would be vertically constant. If, however, the vertical profiles of model clouds (or real clouds) are different from these assumptions, using only model cloud-top CDNC or trying to replicate the retrieval's vertical weighting may not be the most suitable approach. Further work would be needed to investigate this, but such an analysis would be complex and beyond the scope of this study.

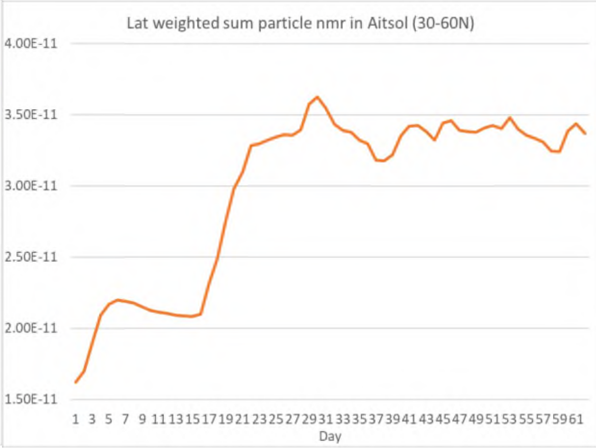


Minor Remarks:

• P5, L180-187: The authors first perform a 60-day spin-up for the aerosol fields. Are these global simulations nudged to the actual meteorology?

The global meteorology is initialised every 24 hours from Met Office analyses data. We expect that this has a similar effect to nudging although the relaxation time is usually shorter in nudging (typically 6 hours). We initially used 30 day spin-up period but found a drift in the simulation and hence made it longer. By the end of 60 day spin-up, the simulation appeared to have reached steady state as shown below (sulphate mass mixing ratio (left) and particle number mixing ratio (right) in Aitken mode in 30-60N).





• P7, L244: *"... Sec'on 2.9 ..."* This should be 2.8.

We thank the referee for pointing out the error. It has been corrected.

• P18, L425: *"... 1.57 ..."* It is 1.56 in Fig. 8.

We appreciate the referee's careful reading and have corrected this error.

• P21, L476: *"... enhancement for volcano on versus volcano off in the first week in the plume region ..."* If I understood correctly, this should be the ERUPTION effect, so I would call it like that.

Thank you for bringing this to our attention. We have added '(ERUPTION effect)' in the text. Likewise, "out of plume to in-plume values" should be the TOTAL effect. Also the ERUPTION effect in the first week should actually be ~20% enhancement instead of ~10%. So the text has become like this;
"The UM liquid water path remains largely unchanged (except the third week) when comparing out of plume to in-plume values (TOTAL effect) but does show a ~20% enhancement for volcano on versus volcano off (ERUPTION effect) in the first week in the plume region."

• P21, L482: *"... the in-plume and out-of-plume regions ..."* If I understood correctly, this should be the LOCATION effect.

The referee is correct in that the comparison between in-plume and out-of-plume regions will give us the LOCATION effect. However, this part of the text talks about the differences in the meteorological environments in two regions. We consider it as the cause of the LOCATION effect and so would not call it the LOCATION effect itself.

References

Haghighatnasab, M., Kretzschmar, J., Block, K., and Quaas, J.: Impact of Holuhraun volcano aerosols on clouds in cloud-system-resolving simulations, Atmos. Chem. Phys., 22, 8457–8472, https://doi.org/10.5194/acp-22-8457-2022, 2022.

Peace, A. H., Chen, Y., Jordan, G., Partridge, D. G., Malavelle, F., Duncan, E., and Haywood, J. M.: In-plume and out-of-plume analysis of aerosol–cloud interactions derived from the 2014–2015 Holuhraun volcanic eruption, Atmos. Chem. Phys., 24, 9533–9553, https://doi.org/10.5194/acp-24-9533-2024, 2024.