Review of "In-plume and out-of-plume analysis of aerosol-cloud interactions derived from the 2014-15 Holuhraun volcanic eruption" by Amy H. Peace et al.

This manuscript assesses the impact of the 2014-15 Holuhraun volcanic eruption on cloud properties over the region for the month of September and evaluated the influence of varying meteorological conditions at synoptic scale on aerosol-cloud interactions. The authors took a plume analysis approach, in which the aerosol effect is evaluated as the contrast between in- and out-of- plume, using both satellite observations and UKESM1 simulations. This approach has been previously used to study ACI associated with the 2014-15 Holuhraun event, but the authors stated that they upgraded the plume-masking method and extended the analysis time period. They found statistically significant cloud microphysical changes (increase in N_d and decrease in r_{eff}) during the first two weeks, both in observations and in simulations. A statistically significant shift in LWP distribution is observed but not modelled in the first two weeks. A synoptic driven shift in airmass origin is observed in week-3 and attributed to the lack of observation and simulation of aerosol effects during week-3.

The manuscript is well written and easy to follow in general, but I did find some places where inherit assumptions and justifications need to be clarified. This work is intriguing in many ways and the topic is of great interest to others in the community, but I think the authors should consider/address these major points (listed below) first to make the conclusions publishable.

Major comments:

- A major question I have after reading this work is that where is the cloud relative to the volcanic plume, is the vertical distribution of SO_2 (and its relative location to the cloud layer) changing with time, and how does this affect the cloud property changes you observed/modelled and the interpretation of them? To be more specific…
    - What is eruption (injection) altitude? And what is the typical cloud top height in this region?
    - Are plumes always in contact with the clouds? Or there are times when they are separated?
    - How good is your assumption of equally distributed SO_2 between 0.8 to 3 km in your ESM setup?
- Regarding the results from UKESM simulations:
    - When interpreting/comparing modelled results to observed results, how do you address the fact that LWP adjustment due to aerosol perturbations is uni-directional by design in ESMs (i.e., only precip-suppression is parameterized, and entrainment-feedbacks are not represented), meaning modelled increase in LWP is likely exaggerated.
    - What, physically, can we learn from comparing LWP in- and out-of- plume in the UKESM-Hol simulation when we know it's not representing the full chain of LWP responses?
    - Have the authors considered running ensemble simulations to relax the assumption on uniformly distributed SO_2 profile by varying it to create ensemble members?
    - When I compare cloud properties in-plume between Hol and Ctrl simulations, I see much stronger changes (which we know is causal aerosol effect) compared to the results you get from the in- vs out- plume method; how do you reconcile this

  - difference? Does this mean the in vs out method is still heavily confounded by meteorological covariations?
  - What's the interpretation of the large difference in out-of-plume cloud properties between -Hol and -Ctrl? Aerosol effects or meteorological difference between simulations?
- I wonder as the eruption goes on, do you see a dilution effect of the SO_2 plume, such that the background (out-of-plume) is getting more polluted with time? Does this contribute to you week-3 lack of signal in N_d? I wonder if you could show histograms of SO_2 comparing in- and out-of- plume, similarly to Fig. 3?
- Your bounding box size varies from day-to-day, meaning the degree of meteorological confounding effect also varies from day-to-day in your analysis, how do you address this issue when you compare cloud changes among days and group them into weeks? Also, since some boxes cover land, do you screen out land clouds? I think you only mention this in the caption of one of the figures, I would bring it up clearly in the methodology.
- I had hard time wrapping my head around the LWP responses, particularly when the statistical testing method disagrees with the mean changes between in- and out-. I understand the source of discrepancy, but I am a bit concerned about a lack of high-level take-away of these results in the context of ACI, i.e., should we take away with the message that LWP response is weak and its sign is undiscernible? Personally, I prefer the statistical testing method that focuses on the distribution shift rather than a change in the mean, which could be driven by outliners and not representing a physical response.
- Regarding the satellite retrieved cloud properties,
  - Weren't you concerned about getting mixed-phase or ice clouds when your 1-5km cloud-top constraint is well above the freezing level?
  - For N_d retrievals, high-SZA and low-CF have been shown to produce unreliable retrievals (e.g., Grosvenor et al. 2018), I would put extra constraints on these two variables.
  - Does SO_2 plume in the scene affect satellite microphysical (tau and r_{eff}) retrievals?

Minor comments:

- What's the main reason and benefit for extending analysis to 4 weeks? Are you targeted to capture temporal evolution in ACI and/or its timescale? Or just want to explore the influence of different synoptic/meteorological patterns? If it's the latter, why only one month is analyzed when you can do this for the whole eruption time period? Why group your analysis by week (is there any physical reason? Timescale assumption? Any major synoptic pattern shift at weekly scale)? I think these inherit assumptions and justifications need to be layout upfront clearly.
- Information on UKESM initialization and boundary conditions needs to be added.
- Need to mention why there are missing days in the analysis (e.g., Fig. 4).
- Line 87, check sentence "… to and …"
- Figures, missing lat/lon labels when maps are shown, and please add color code to the caption.
- Line 213, "e.g." should be inside the parenthesis.
- Line 226, I think Sep-11 and Sep-25 in Fig. 2 are examples where they do not agree, please discuss and reword.

- Line 231, you meant “north” of the domain? as “top” refers to the vertical direction.
- Line 270-271, after reading through the whole manuscript, I am still missing an explanation on week-4 LWP responses, observed and modelled.
- I feel Fig. S2 is worthy of being included as a main figure.
- Line 303, what is the “reason” that you’re referring to? And how do you know it’s necessarily the same reason in the simulation, rather than different reasons leading to the same results? perhaps need to reconstruct this sentence.
- Table 1, cloud fraction responses in simulations? Line 328, perhaps better to define perturbation in the main text and use a mathematical expression.
- Line 344-347, this discussion in not very clear and hard to follow, especially the use of terms like ‘updraft-limited’ and ‘aerosol-limited’, which are not straightforward concepts and need to be explained/introduced.
- Line 365-371, why choose modelled meteorological conditions instead of ERA reanalysis? And define how is LTS calculated.
- Lines 412-414, any speculation on why this is the case?
- Lines 484-486, I think this needs to be mentioned upfront, I have been wondering about CF responses when I read the results.

References

Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., Boers, R., Cairns, B., Chiu, J. C., Christensen, M., Deneke, H., Diamond, M., Feingold, G., Fridlind, A., HÃŒnerbein, A., Knist, C., Kollias, P., Marshak, A., McCoy, D., Merk, D., Painemal, D., Rausch, J., Rosen- feld, D., Russchenberg, H., Seifert, P., Sinclair, K., Stier, P., van Diedenhoven, B., Wendisch, M., Werner, F., Wood, R., Zhang, Z., and Quaas, J.: Remote Sensing of Droplet Number Con- centration in Warm Clouds: A Review of the Current State of Knowledge and Perspectives, Rev. Geophys., 56, 409–453, https://doi.org/10.1029/2017RG000593, 2018.