Dear Dr Folch,
We thank you a lot for your valuable comments and suggestions. We addressed them as explained below.
The reviewer's comments are repeated in **bold letters**, our replies are given in standard font, and text modified or added to the manuscript is given in blue.

This paper uses the ICON-ART modelling system to study the effects of volcanic aerosol dynamics (alterations in aerosol size and composition due to particle aging) and aerosol-radiation feedbacks on the dynamics of volcanic clouds. It is known that the strong absorption of fine ash particles can cause thermal disequilibrium with the surrounding atmosphere, potentially altering the atmospheric dynamics. However, in-depth studies are scarce in the literature and this paper is an important step forward. The authors show results for the 2019 Raikoke eruption, using measurements from different satellite instrumentation for model validation; TROPOMI and AHI for SO2/ash column mass retrievals, and MOIS/VIIRS/CALIOP/OMPS-LP for cloud top height. It is difficult to extract conclusions from a single example but, overall, I think this paper is very relevant to show the potential effects of both phenomena on model forecasts. I do recommend publication with minor revisions detailed below.

Thank you very much for the insightful review. Your comments and questions helped us a lot to improve the manuscript.

1. ICON-ART is run for 3 scenarios: AERODYN_rad (aerosol dynamics + radiation), no_AERODYN_rad (no dynamics) and AERODYN_no_rad (no radiation), which allow isolating the effects of dynamics and radiation. These are actually in competition, with dynamics enhancing premature settling and radiation uplifting the cloud (as nicely shown in Figure 8). To what extent can these two effects counterbalance? This is somehow discussed in Sec 3.3., but it would be great to compare AERODYN-rad results with the no_AERODYN_no_rad ICON case. Note also that, to my knowledge, all operational volcanic cloud forecast systems do not include neither dynamics nor radiation and therefore the no_AERODYN_no_rad (not shown) would actually mimic current setups.

We agree that operational volcanic cloud forecast centers do neither include dynamics nor radiation interaction. Therefore, a comparison with this simulation case would indeed be beneficial. As we ran ICON-ART in the setup no_AERODYN-no_rad, we add some of these results to the manuscript.

Updated Table 2:
We include the no_AERODYN-no_rad scenario in Table 2.

We add to l. 279:
The fourth scenario represents the status quo of operational volcanic cloud forecast. It considers neither aerosol dynamic effects nor aerosol-radiation interaction.

Updated Fig. 4:
We replace the AERODYN-no_rad by the no_AERODYN-no_rad scenario. For details, please refer to answer of comment 2.

Updated Fig. 6:
We include the two no_rad simulation scenarios in Fig. 6. Furthermore, additional dates with a comparison between CALIOP and ICON-ART model results are displayed in the Appendix of the manuscript.
We add to l. 358:

A similar conclusion can be derived from the AERODYN-no_rad and no_AERODYN-no_rad scenarios in Fig. 6 (e) and (f), respectively. Although, both are missing the most prominent feature between 49° N and 51° N at around 16 km, they show the same behavior in terms of aerosol dynamic effects.
Additional dates of CALIPSO measurements are displayed in Appendix A.

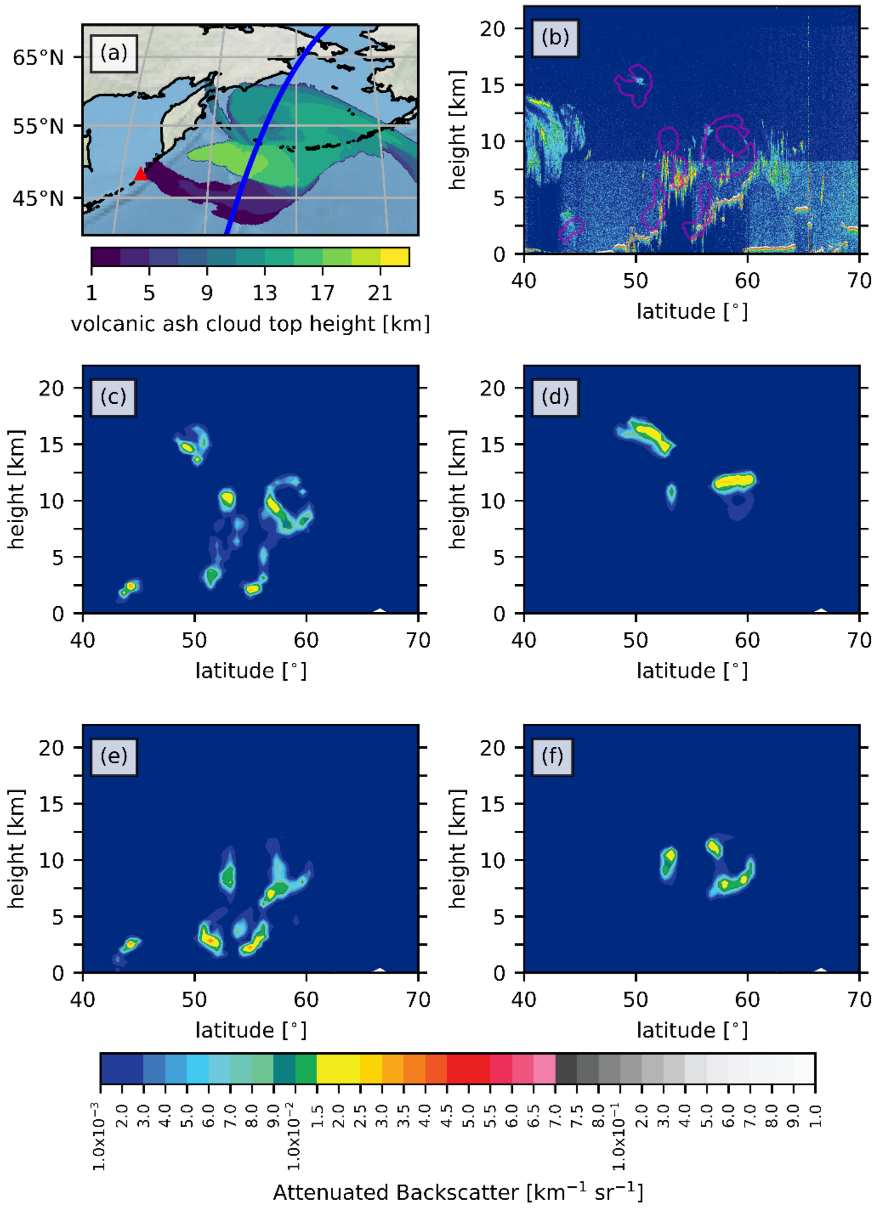


Figure 6. (a) CALIPSO ground track on 23 June 2019, around 15:00 UTC in blue color and location of Raikoke volcano as red triangle. The contour map shows the volcanic ash cloud top height for the AERODYN-rad scenario. (b) The CALIOP attenuated backscatter for 532 nm for the satellite position between 40° N and 70° N is displayed in the top right panel. The magenta line shows the 0.002 km^{-1}sr^{-1} contour of AERODYN-rad at 15:00 UTC. Middle and lower panels: Total attenuated backscatter for 532 nm of volcanic aerosols under the CALIPSO ground track on 23 June 2019, for the 15:00 UTC model output are displayed. (c) shows the result for AERODYN-rad, (d) for no_AERODYN-rad, (e) for AERODYN-no_rad, and (f) for no_AERODYN-no_rad, respectively.

Updated Fig. 8:
We include the no_AERODYN-no_rad scenario plume top height in Fig. 8. Furthermore, we add an error bar for the OMPS measurement in the same figure (as requested by referee #1).
For further explanation we rephrase l. 410:
A distinct difference prevails between the two scenarios with radiative interaction (yellow and green curve) and the two without radiative interaction (pink and orange curve).

And add to l. 420:
As for accumulation mode particles, in the no_AERODYN-no_rad scenario (orange curve) coarse mode particles also tend to stay on the same height level.

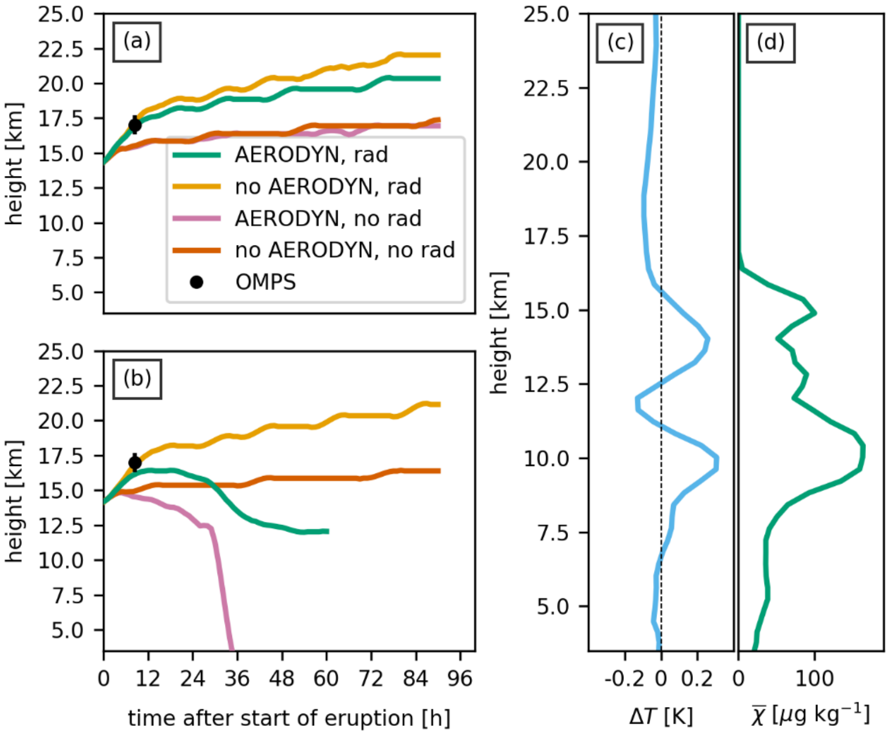


Figure 8. (a) and (b) Evolution of height of volcanic ash cloud top after the onset of the eruption on 21 June 2019, at 18:00 UTC. The yellow curve represents the no_AERODYN-rad scenario, the green curve AERODYN-rad, the pink one AERODYN-no_rad, and the orange one represents the no_AERODYN-no_rad scenario. Panel (a) shows the ash cloud top of particles in the accumulation mode, (b) of particles in the coarse mode, respectively. The black circle depicts the volcanic cloud top height obtained from OMPS-LP. (c) Mean temperature difference (AERODYN-rad – AERODYN-no_rad) in volcanic ash cloud columns on 23 June 2019, 12:00 UTC. (d) Mean volcanic ash concentration $\overline{\chi}$ for the same model columns as in (c) for AERODYN-rad.

2. Figure 4 is very interesting but panels (c)-(e) (and (d)-(f)) are difficult to distinguish and should highlight differences better (e.g. using a log scale). A better option could be plotting relative differences (in percent) between both model configurations, using AERODYN_rad as the “true”. Is it a 10\% or a 100\%? Difficult to say from (d)-(f) with the contour range used.

We updated Fig. 4 in two ways. First of all, we replaced panels (e) and (f) by the total column ash mass loading of the no_AERODYN-no_rad case. Secondly, we added panels (g) and (h) which are showing the absolute difference between the two simulation scenarios AERODYN-rad – no_AERODYN-no_rad.

We add to and rephrase l. 293:
These differences are mainly restricted to the slightly higher mass loading in panel (e) and small differences in the volcanic cloud structure. For the first day after the eruption, the aerosol dynamic effects and the aerosol-radiation interaction have only a minor influence on the volcanic ash mass loading.
We add to and rephrase l. 300ff.:
Compared to these two simulations, the averaged AHI measurements (Fig. 4 (b)) show values for the maximum ash mass loading that lie in between the two simulation scenarios. In panels (g) and (h) the differences between the two are highlighted by the absolute difference of AERODYN-rad – no_AERODYN-no_rad. It shows that considering aerosol dynamics and aerosol-radiation interaction results in lower volcanic ash mass loadings in most parts of the volcanic cloud. Only for the first day after the eruption, the volcanic cloud seems to be shifted slightly north in the AERODYN-rad scenario compared to the no_AERODYN-no_rad scenario, as the difference plot shows some positive values between 160 – 170° N.

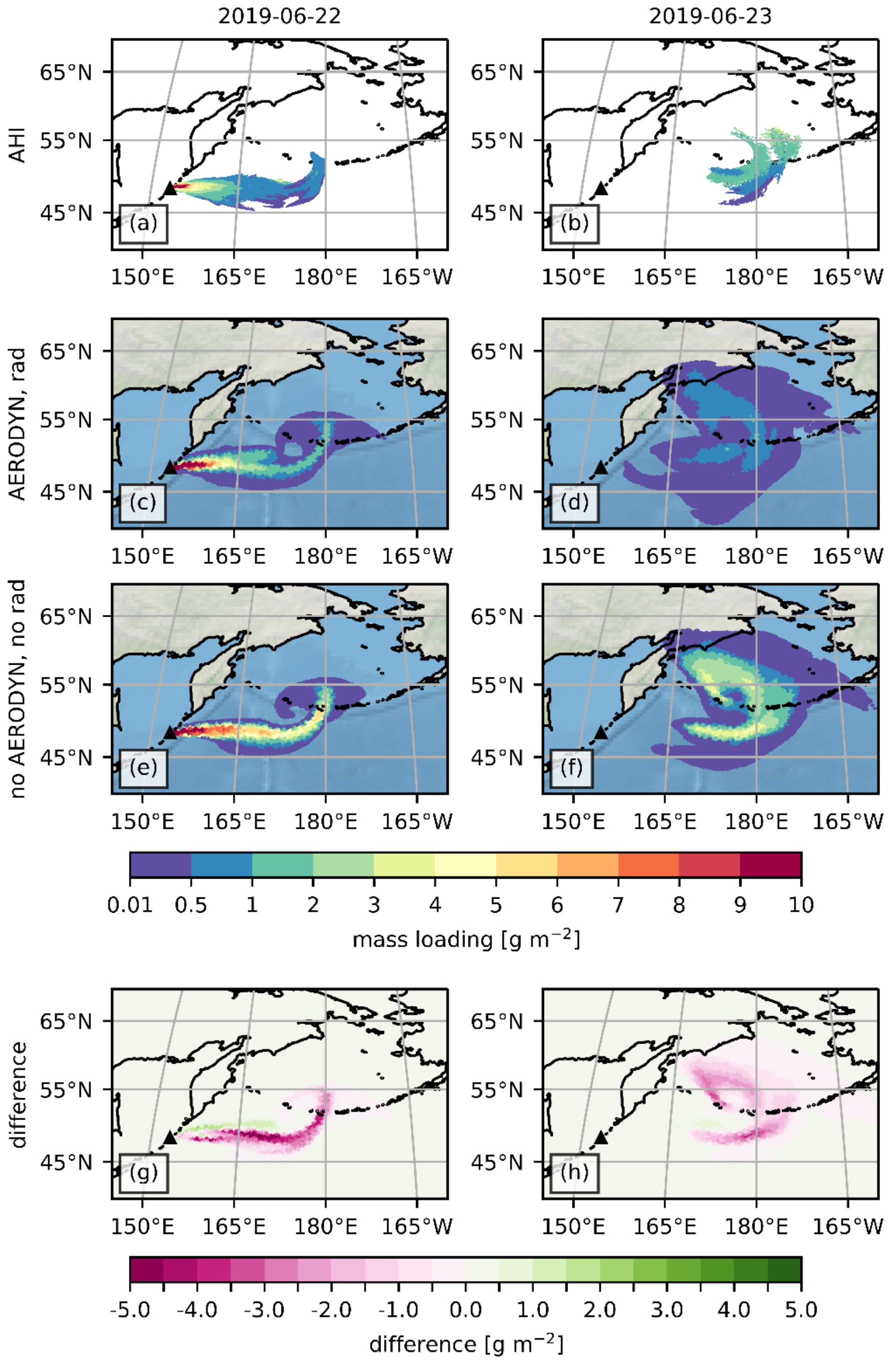


Figure 4. Daily mean total column mass loading of volcanic ash on 22 June (left column) and 23 June 2019 (right column). Top row (panel (a) and (b)) shows results measured by AHI on-board Himawari-8. The middle and lower row (panel (c) - (f)) show ICON-ART results for AERODYN-rad and no_AERODYN-no_rad, respectively. The black triangle depicts the location of Raikoke volcano. Panels (g) and (h) show the absolute difference between the two simulation scenarios.

3. On the other hand, and related to the point above, I missed some figure or text showing the impact on the atmospheric dynamics when switching on the AERODYN_rad module. To what extent is the vertical wind field advecting the cloud modified by thermal perturbations? Can you quantify? I understand that this question may fall beyond the objective of the paper, but it could be of interest to the volcanic cloud modelling community. Ensemble forecast strategies are gaining more and more attention, and these rely on perturbing uncertain variables like the eruption source parameters or the wind field (but rarely the vertical component). As a result, an interesting question it to assess whether (vertical) wind perturbations caused by radiation feedbacks are comparable to typical uncertainties in NWP models. If in the range, an ensemble of offline models could still capture this effect, at least to some extent.

If we look at the most pronounced lifting of the volcanic cloud top height (approx. 3 km, compare Fig. 8) during the first 12 h of the simulation, we obtain an average vertical lifting velocity of 0.07 m/s. This lifting is only visible for simulation scenarios with radiation interaction.
We determined the vertical velocity difference between the AERODYN-rad and the no_AERODYN-no_rad scenario as well as between the AERODYN-rad and AERODYN-no_rad scenario. Both comparisons show comparable numbers. For the comparison, we only consider grid cells which are within a vertical column which contains a volcanic ash mass loading > 0.01 g m^{-2}. The maximum absolute difference that appears locally during the first 12 h of the eruption is in the order of 0.19 m/s with a 98th percentile of 0.05 m/s. We would like to note, that these vertical velocity perturbations strongly depend on the spatial resolution. For a finer resolution, locally we would expect higher vertical velocities.

We include this information in the manuscript in l. 431ff.:
The resulting vertical velocity perturbation Δw is in the order of 0.1 m s^{-1}. For this purpose, we analyzed the difference in vertical velocity between the AERODYN-rad and AERODYN-no_rad scenario during the first 12 h after the eruption. Only grid cells in model columns which contain a volcanic ash mass loading > 0.01 g m^{-2} in both scenarios are considered. Locally, Δw reaches 0.19 m s^{-1} with a 98th percentile of 0.05 m s^{-1}. This agrees well with the vertical lifting of the volcanic cloud top height of around 3 km during the first 12 h ($\overline{w}$ = 0.07 m s^{-1}).

4. The aerosol dynamics module (ARODYN) has pre-defined initial aerosol size distributions, which (if I am not wrong) are evolved according to prognostic equations. How does the aging mechanism depend on this initial condition? Particle distributions can vary notably from one eruption to another, and a single representation could be misleading.

We agree that particle size distributions (PSDs) can vary notably from one eruption to another. For the Raikoke simulation, we defined the emitted PSD as specified in Table 1 in the manuscript. This emitted PSD changes over time as particles age and sediment. Very often, there is lack of direct measurements when it comes to the PSD of volcanic ash from one particular volcanic eruption like Raikoke. To overcome this limitation, we used the PSD data from five eruptions (listed in the following table) to calculate a generic PSD for volcanic ash in ICON-ART (shown in the following figure). This PSD captures the variability of fine and coarse particles. We are aware of the uncertainties associated with

this generic PSD. Nevertheless, even direct measurements are subjected to large uncertainties and might fail to represent the variability of the PSD.
The aging mechanism which is implemented in AERODYN depends on the PSD. As condensation of gaseous species on existing particles, coagulation, sedimentation, and deposition directly depend on the particle diameter. However, it needs further studies in order to quantify how the aging mechanism depends on the emitted size distribution.

Table: Volcanic eruption for which validated ash PSD data exist (from http://www.ct.ingv.it/iavcei/results.htm)

Eruption	**Montserrat (West Indies)** **31 March 1997**	**Mt. St. Helens (USA)** **18 May 1980**	**Ruapehu (New Zealand)** **17 June 1996**	**Spurr (Alaska)** **16-17 September 1992**	**Eyjafjallajökull (Iceland)** **14 April-21 May 2010** **(data for 4-8 May 2010)**
Eruption type	dome collapse (co-PF plumes)	Plinian+coignimbrite (strong plume)	sub-plinian (weak plume)	sub-plinian	long-lasting weak plume

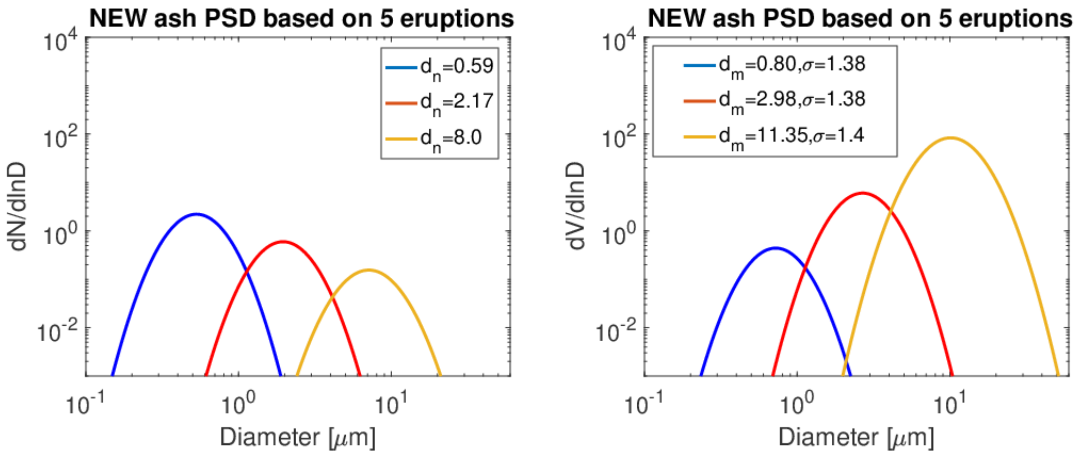


Figure: Calculated ash PSD based on the data available in literature

Bonadonna, C. and Scollo, S.: IAVCEI Commission on Tephra Hazard Modelling, http://www.ct.ingv.it/iavcei/results.htm, last access: 03 September 2020, 2013.

We add to l. 273:
They are based on data from Bonadonna and Scollo (2013).

5. Model validation. Several plots compare model results with observations. However, I missed some quantitative metric values; e.g. SAL, Figure Merit of Space or others. These are by far more objective than color plots (e.g. Figs 4, 5), which can trick depending on the scale and color binning. Given that a main objective of the paper is to “assess if representations of aerosol dynamics and aerosol-radiation interactions are beneficial for forecasts”, quantitative metrics would help asking this question more objectively.

We apply the SAL method following Wernli et al. (2008) in order to compare the total column volcanic ash mass loading AHI retrieval with our model result.

Wernli, H., M. Paulat, M. Hagen, and C. Frei, 2008: SAL—A Novel Quality Measure for the Verification of Quantitative Precipitation Forecasts. Mon. Wea. Rev., 136, 4470–4487, https://doi.org/10.1175/2008MWR2415.1

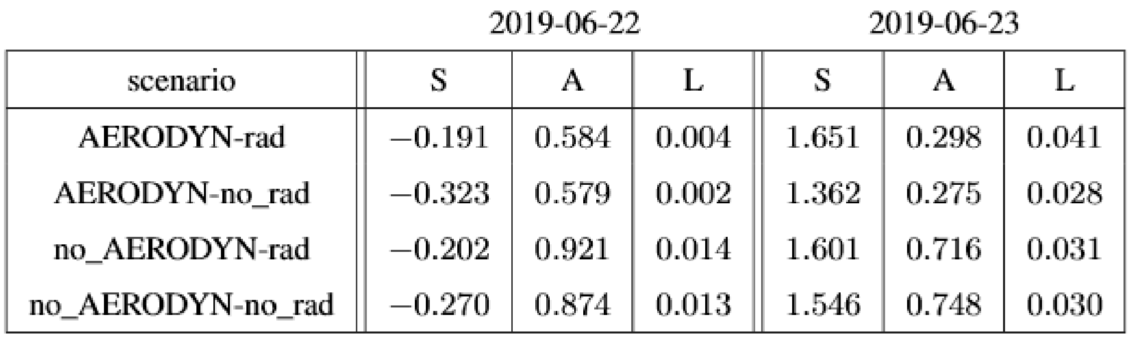
Wernli, H., C. Hofmann, and M. Zimmer, 2009: Spatial Forecast Verification Methods Intercomparison Project: Application of the SAL Technique. Wea. Forecasting, 24, 1472–1484, https://doi.org/10.1175/2009WAF2222271.1

We add the following paragraph to the manuscript in l. 302ff.:

In order to compare our ICON-ART results in an objective manner with the AHI observations, we make use of the SAL method. This quality measure has been introduced by Wernli et al. (2008) and has been extensively discussed by Wernli et al. (2009). The method identifies objects in a 2D field (e.g., total ash mass loading) and quantifies the differences between model and observation in structure (S), amplitude (A), and location (L). A value of 0 implies perfect agreement. We apply the SAL method with a fix threshold value to identify objects R* = 0.01 g m^{-2}. The results for the comparison of daily mean total column mass loading between the AHI retrieval and the ICON-ART results are summarized in Table 3.

The location of the volcanic cloud agrees very well with the observation for all dates in all simulation scenarios. The structure of the volcanic cloud shows larger differences compared to observations, especially on 23 June. However, the values are rather similar for the different simulation scenarios. Only the amplitude values differ distinctly among the different scenarios. Simulations with AERODYN are closer to the observation than simulations without aerosol dynamics.

Table 3: Comparison of daily mean total column mass loading of volcanic ash between AHI and ICON-ART results using the SAL method by Wernli et al. (2008).



	2019-06-22			2019-06-23		
scenario	S	A	L	S	A	L
AERODYN-rad	−0.191	0.584	0.004	1.651	0.298	0.041
AERODYN-no_rad	−0.323	0.579	0.002	1.362	0.275	0.028
no_AERODYN-rad	−0.202	0.921	0.014	1.601	0.716	0.031
no_AERODYN-no_rad	−0.270	0.874	0.013	1.546	0.748	0.030

6. Line 84. “density values less”?

We agree that this formulation is a bit misleading and hope that the reformulation makes it easier to understand.

We change the sentence on p.3 l.84 from:

Only data with the quality descriptor ’qa_value’ larger than 0.5 and total vertical column density values less than 1000 mol m^{-2} were used.

to:
Only data with a quality value larger than 0.5 (as recommended in the TROPOMI product user manual) and total vertical column density with values less than 1000 mol m^{-2} were used.

7. Line 257. It is stated that the source term in ICON-ART is set between 8 and 14 km a.s.l. Does it mean a 6 km thick cloud? This seems quite inconsistent with the TROPOMI retrievals, which assume 1 km thickness at 15 km a.s.l.

Yes, in the model simulation we emit a 6km thick cloud of ash and SO_2. Our emission parametrization for ash and SO_2 is based on satellite observations (as well as results of Plumeria and FPlume). The configuration of the emission height has been done specifically for the Raikoke eruption in 2019 and is based on satellite observations and volcanic monitoring reports (Sennet, 2019).
Whereas, the TROPOMI retrieval assumptions have been set for a much broader range of scenarios. The retrieval algorithm can be run with one of four different assumptions on where the SO_2 is located in the atmosphere. This could either be a vertical profile modeled by the global chemistry transport model TM5 or a 1 km thick box in either 1 km, 7 km or 15 km. Comparisons with other satellite products showed, that the assumption of a 1 km box in 15 km a.s.l. gave best results, although, the retrieval assumption does not match with the actual SO_2 distribution in the atmosphere after the Raikoke eruption.