

Review of manuscript “Improved estimation of volcanic SO₂ injections from satellite observations and Lagrangian transport simulations: the 2019 Raikoke eruption” by Cai et al.

General comments:

The study by Cai et al. combines satellite retrievals from TROPOMI and AIRS with the Lagrangian transport model MPTRAC to give a detailed estimate of the volcanic SO₂ injections for the 2019 Raikoke eruption. By applying an inverse modelling technique, the authors give a detailed estimate of the time- and altitude-resolved SO₂ emission for this eruption. This study investigated a range of conditions in the latest version of MPTRAC (e.g., using various SO₂ retrievals, varying diffusion parameters, scaling of the initial SO₂ mass, including a new OH chemistry module), which results in a detailed consideration of the estimated SO₂ injections.

Initialising forward trajectories with this new SO₂ estimates for Raikoke can reproduce the SO₂ mass variations and spatial dispersion of the volcanic cloud retrieved by TROPOMI during the first 10 days after the eruption. Finally, the impact of diffusion is investigated, showing that it is too strong and as a result the model cannot capture the internal structure of the simulated SO₂ cloud well.

This work is very relevant to the atmospheric modelling community. Detailed eruption source parameters for volcanic eruptions are known to be difficult to determine. This study is a useful addition to the existing literature demonstrating that inverse modelling techniques are very useful to create more detailed eruption source parameters to initialise model simulations.

The manuscript is rather long, but I cannot see how to make it much shorter other than merging some of the figures. It is well written, and the figures are of a good quality and the authors give clear interpretations of the data. I have some minor concerns about some of the methodology used in this study, as outlined in the specific comments below. However, the overall work presented in the manuscript is good. I therefore recommend minor revisions to address the points outlined below before publication.

Specific comments:

L11: Satellites do not directly observe SO₂, but (as explained in sections 2.1 and 2.2) uses various bands in the infrared/UV spectra to retrieve estimates for SO₂. Therefore, in general it is better to use the terms ‘retrievals/retrieved’, rather than ‘observations/observed’ when discussing the satellite products. Please check carefully throughout the manuscript.

L114: “The AIRS... upper levels.” I think this sentence needs a reference that supports this statement.

L115: ‘particularly carefully’. Please avoid non-scientific terms. It should be clarified how the authors used the daytime data.

L130: In this study the results from the 15 km retrieval are presented. However, based on Fig. 3, you could argue that the assumed SO₂ height for the 15 km retrieval product is too high in the atmosphere for this eruption and that the 7 km retrieval from TROPOMI could be considered equally realistic (especially for the second and third phase). Did the authors investigate the impact of using the 7 km retrieval on their results? The 15 km and 7 km

retrieval products for TROPOMI can result in different SO₂ column mass estimates, which would in turn could also influence your estimate of the total emitted mass. It would be very interesting to understand if you would still get the reported 2.1 ± 0.2 Tg estimate when applying the 7 km retrieval. I think some discussion on this potential source of uncertainty should be included in the paper.

L155: "... reaction with the hydroxyl radical (OH)." What data is used to provide the OH field for the MPTRAC simulations?

L179-190: The work presented here samples trajectories from each satellite footprint between 0-25 km and combine them to obtain a best estimate of the SO₂ emission. However, how certain is it that there is a single best solution to the problem? Is it possible that a different emission profile not found by this method could give an equally good comparison with the TROPOMI retrievals?

For example, assume two hypothetical layers (say at 2 km and 7 km) at a given TROPOMI pixel location that both can be retraced to the volcano. I am not sure I understand how the backward trajectory method is able to determine which layer contained the SO₂ mass (or is able to determine the ratio between the two layers)? If I understand correctly, based on the sensitivity of the TROPOMI satellite (fig. 1), the method seems to be biased towards the higher altitudes, as more trajectories are released from the higher layers (7 km in our example case). But in our example, it is also possible that the mass was emitted in the lower layer, as it would give the same column total mass in TROPOMI. Is the backward trajectory method able exclude either possibility?

L193: "fixed e-folding lifetime". How realistic is it to use a constant e-folding lifetime for the entire altitude range considered? As mentioned by the authors (L.296-302, Fig.10), the lifetime of SO₂ varies strongly between the troposphere and the stratosphere. How should we interpret this fixed e-folding lifetime? Is it an altitude-weighted average lifetime for the SO₂?

L199: "To achieve a total injection of 2.1 Tg." So far, no evidence is given why 2.1 Tg would be more realistic. I think the authors should point to the discussion presented in section 3.1.2 or give a short explanation here on why the results have been tuned to 2.1 Tg.

Fig.3: What does the black line in panel 3a represent?

L202: "...prominent second and third plume..." Looking at the video's provided in the study by de Leeuw et al, their dispersion simulations show that part of the first plume is dispersed back to the location of the Raikoke volcano around the same time this study identified the third plume (27-28 June). This is also shown by the red trajectory in figure 11g. How would the backward trajectory methodology deal with such an event and is it possible that this third peak in the shown analysis is partly a reoccurrence of the first plume at the Raikoke location. If back trajectories hit the source location multiple times, would it pick the first hit only, identifying a 'new' source, or would it consider the possibility of multiple overpasses over the volcano at earlier times?

Related to this point, the video for SO₂ dispersion provided in the study by de Leeuw et al. also shows that part of the SO₂ cloud (using the VolRes profile) initially moves into a North-West direction, followed by it moving back and across the Raikoke volcano location around the 25th of June. Based Fig.11 in the current paper, this part of the plume is not present in the

presented dispersion experiment (panel 11b), while it is visible on the TROPOMI retrieval (panel 11a). This suggests that it might be linked to emissions at low altitudes (<5km). I wonder if part of the 2nd plume in fig. 3 could be related to this residual of the low altitude first plume that moves across the Raikoke volcano during the initial few days after the eruption and is misinterpreted as an additional emission of SO₂ at higher altitudes (see also point L179)?

I think it could be a great addition to the paper if it is possible to initialise the MPTRAC model using the VolRes profile as input and compare it with the results presented in this study (e.g. figs. 9, 11, 15).

L203: How is the tropopause defined?

L248: "...it matters how many days of satellite observations are used...". How many days are used for the results presented in this study (e.g. figure 3)? I think this should be specified in the manuscript. I also think some additional explanation is missing that describes how multiple days of the TROPOMI satellite retrievals are combined to reconstruct the SO₂ injections. When using 12 days of retrievals, does this mean that all the earlier overpasses are still considered? Or does this mean that the back-trajectories are calculated for 12 days to reconstruct the signal for this specific overpass?

Figure 9: Would it be possible to extent the figure to longer timescales? Based on figures 3, 10 and 13-15 the data is available, so I wonder why it wasn't included here?

L270: Why did the authors choose to implement the constant injection rate to represent 1.5 Tg and not 2.1 Tg like the other simulations? Earlier it is established that 1.5 Tg does not give realistic values, so is it considered for the constant injection rate. When looking at figure 9, moving up the constant emission simulation by 0.6 Tg, apart from the initial peak, you get a much better comparison with TROPOMI and the results for this simulation also fall within the satellite uncertainty range after several days, like the other two simulations.

Also, it is not clear to me if the chemistry module is used for the constant injection rate simulation or whether it uses the 14-day exponential decay (similar to the orange line in figure 9). Based on how smooth the removal is in figure 9 for the purple line, I think the latter is true, but this needs to be clarified.

L277: Which chemical reactions are included in the OH chemistry module and what are the reaction rates? I think a description of the chemistry module or a citation explaining the module should be included.

L283: "... are consistent with the total SO₂ mass derived from TROPOMI estimations." If I interpret figure 8 correctly, it seems that the MPTRAC data in figure 9a are initialised using the TROPOMI retrievals during the first 10-12 days. Therefore, I wonder how independent the two datasets are for the period shown in the figure and whether the very good agreement is a direct consequence of the way how the MPTRAC simulations are initialised using the same retrievals to which it is now again compared against.

L300-305: The constant injection rate simulation assumes that the mass is emitted uniformly between 5 and 15 km altitude. Assuming a tropopause at 10 km, this means that approximately 50% of the mass is emitted into the troposphere, where the lifetime of SO₂ is

much shorter (as seen in figure 10). If you would use a constant injection rate that has a more realistic profile with more of the mass emitted into the lower stratosphere/upper troposphere, how much would this improve the comparison?

L327: "... mainly associated with transport of SO₂ in the lower troposphere (between 0 and 5km), which was not represented in both initializations". Why did the backward trajectory method not manage to track back the TROPOMI footprints to the volcano for this part of the plume? Based on figures 11a and b, most of the southern branch retrieved by TROPOMI is not present in the MPTRAC simulations. The fact that the method does not seem to capture this large area of the plume associated with the lower emission altitudes makes me doubt the robustness of the method (see point L179), especially for lower altitude eruptions? I think this potential limitation should be discussed in the discussion section.

L333: Would it be possible that the part not represented by the constant emission injection rate simulation in panel 11f is linked to emissions at lower altitudes that were also missing in panels 11b-c? I think it would be very useful to repeat the analysis for the VolRes profile (which includes the lower-level emissions) to see if this could explain part of the differences observed.

Fig. 13-14: I think these figures can be combined to one figure with 6 panels, rather than having two separate figures. Label of Fig.13: (CSI, a) -> (CSI, c)

Fig. 15: 'Color coding indicates the column density threshold...' This is incorrect, as the different colors show the different simulation initialisations at a constant detection threshold (5 DU).

L410-414: I think the authors should include the altitudes used in MPTRAC for this part of the analysis. Maybe in a table in the supplementary material could be a good option, as the current manuscript is pretty long already.

L460: I miss a short discussion of the limitations/uncertainties related to the MPTRAC model and the backward trajectory method. One potential impact not discussed is the impact of the lofting of the plume due to the co-existence of ash. In the paper by Muser et al. 2020, a lofting effect was identified for the Raikoke plume during the initial days after the eruption. I can't find any information that the MPTRAC model accounts for this effect and as a result the back-trajectories could be placed at the wrong altitudes in the reconstruction. This in turn could have an impact on the forward simulations on longer timescales.

L474: The study by de Leeuw et al. shows that the 2.0 Tg emission profile overestimates the SO₂ mass from the TROPOMI retrievals during the first few days after the eruption. Therefore, I think this statement would be more accurate when adding 'to match the TROPOMI retrievals on timescales > 1 week'

L475: Are the stratospheric amounts similar for all the simulations considered or is this the value for the most accurate simulation? If increasing the emission to 2.1 Tg for the constant emission case, how much would be emitted into the stratosphere? I think it would be a very useful addition if the authors could include an uncertainty range for the 0.85 Tg estimate using the full range of simulations they have performed.

L483: “From our forward simulations, the second and third plume are potentially overestimated.” Would you be able to identify potential reasons for this overestimation? I think a short discussion what might have caused the overestimation should be included here.

L488 and L499: 2 weeks -> 10 days. In this study only the location and spatial extend during first 10 days are discussed.

L518: I think it would be useful to include the fraction of the mass emitted into the stratosphere (0.85 Tg) here? Especially for climate impact studies, it is mainly the long-term stratospheric part of the plume that is of particular interest, as most of the tropospheric signal will be removed after several days/week.

L519: Better than what? The study does not show how the presented mass estimates compare with mass estimates using the profiles from other studies.

Technical corrections/suggestions:

Order of the figure panels. Some figures have panel b above a (fig. 3) and others have panel b below panel a (fig. 7). Please use one consistent ordering of labelling the panels in the figures to avoid confusion.

L40.: Besides -> Beside

L41: parcels -> parcel

L71: observed from satellite -> retrieved by the satellite

L78: since begin of -> since the beginning of

L85: satellite observations -> satellite retrievals

L106: DU -> Dobson Unit (DU)

L178: As both, AIRS and TROPOMI -> As both AIRS and TROPOMI

L199: turned -> tuned

Fig.7: Differences of -> Differences in

Fig.7: TROPIMI -> TROPOMI

L255: short term -> short-term

L256: long term -> long-term

References: Please check all the references carefully, as some have the DOI included twice (e.g. Hoffman et al 2014). Also use abbreviations for all the journals consistently.