Acp-2020-889
The 2019 Raikoke volcanic eruption: Part 1 Dispersion model simulations and satellite retrievals of volcanic sulfur dioxide

General comments

In this paper NAME simulations of SO2 are compared to SO2 observations from TROPOMI. It is good to see these new observations being used to evaluate the dispersion of SO2 in a quantitative manner using some interesting spatial verification techniques. Generally, NAME appears to perform very well, although the peak SO2 concentrations are underestimated. Overall the paper contains some interesting results, but it is long and contains several repetitions and material that is already in the published literature. The motivation for the experimental design and the hypotheses need to be clearer.

Major comments

1. One aim of the paper is to understand why the observed peak (>20DU) concentrations are underestimated by NAME. There appears to be 3 potential explanations investigated in the paper. (i) The emission profile – too little SO2 emitted into the stratosphere; (ii) the diffusion parameterisation – too much mixing in the stratosphere; (iii) the tropopause height – too high. After reading the paper it was not clear which of these explanations was the dominant factor. I appreciate that it is probably a combination of all 3 but the sensitivity experiments performed seemed a little ad hoc and not best designed to test the 3 hypotheses making it difficult to reach a conclusion. As a result, the abstract contains the results of the sensitivity experiments, but no more general conclusions are/can be included.
2. Given that the focus of the paper is mixing in the stratosphere, it is surprising that there are no references to the extensive literature on this topic. Below are a few suggestions that should be referred to, but there are many more.

Balluch, M. G., and P. H. Haynes (1997), Quantification of lower stratospheric mixing processes using aircraft data, *J. Geophys. Res.*, **102**, 23,487– 23,504
Colette, A., and G. Ancellet (2006), Variability of the tropospheric mixing and of streamer formation and their impact on the lifetime of observed ozone layers, *Geophys. Res. Lett.*, **33**, L09808
Hall, T. M., and D. Waugh (1997), Tracer transport in the tropical stratosphere due to vertical diffusion and horizontal mixing, *Geophys. Res. Lett.*, **24**, 1383– 1386.
Haynes, P., and J. Anglade (1997), The vertical-scale cascade in atmospheric tracers due to large-scale differential advection, *J. Atmos. Sci.*, **54**, 1121– 1136.
Legras, B., I. Pisso, F. Lefèvre, and G. Berthet (2005), Variability of the Lagrangian turbulent diffusion in the lower stratosphere, *Atmos. Chem. Phys.*, **5**, 1605– 1622.

Minor comments

1. Abstract, line 24. I'm concerned that the authors are overstating their result. The paper demonstrates the potential for satellite data to improve the representation of SO2 dispersion for this case study, but I don't think they can claim that it can 'rectify limitations in dispersion models like NAME'. This would require an analysis of many volcanic SO2 clouds and a systematic improvement.
2. Page 3, line 60. The authors state that 'SO2 clouds are potential tracers for the more difficult observable ash clouds', do they mean 'more difficult to observe ash clouds'?
3. Page 3, line 63. Models are plural so 'is' should be 'are' I think.
4. Page 4, line 80. What is the 'unprecedented resolution', can the authors be quantitative?
5. Page 5, line 100. Repetition of information from line 79, page 4.

6. Page 5, line 111. Repetition.
7. Page 5, lines 123 and 124. Why are the uncertainties in the SO2 retrievals different?
8. Page 5, line 129. Why is the SO2 layer assumed to be 15km agl?
9. Page 5, line 131. Why are 15km and 7km chosen specifically for sensitivity testing? Can a more extensive systematic sensitivity test be carried out?
10. Page 6, line 136. Please include a reference to justify the 0.3DU threshold used.
11. Page 6, line 140. NAME is a Lagrangian model so doesn't have grid cells. Are you referring to the output grid on which SO2 concentrations are calculated?
12. Page 6, line 143. What is ash-inference? Do the authors mean ash interference?
13. Page 7, line 176. What is the vertical resolution of the meteorological data at the tropopause height?
14. Page 7, lines 177-187. This textbook information is generic to all Lagrangian models. Does it need to be included in the paper?
15. Page 7, line 191. Sigma is the standard deviation of the velocity not the velocity I think.
16. Page 10, table 2. What does full refer to? K_meso?
17. Page 10, table 2. Why is K_meso changed and not K_turb? I couldn't find any motivation for these experiments. Reference to the published literature on stratospheric mixing may help to motivate this experiment.
18. Page 11, line 260. How is the lower stratosphere defined?
19. Page 11, line 262. Why do you need to perform a separate NAME simulation to increase the emissions? Doesn't the output just scale by +33%? Perhaps I have misunderstood this experiment.
20. Page 11, line 264. Repetition of IASI satellite overpasses.
21. Page 11, line 265-270. The motivation for your StratProfile experiment is not clear to me. What hypothesis are you testing? Are you claiming that the first profiles you used were wrong? If so why?
22. Page 11, line 271 and 275. Comparison of the modelled and observed tropopause height suggests that the modelled height may be too high. Therefore, rather than interpreting the emission profile as emitting too little SO2 into the stratosphere, could an alternative interpretation be that the tropopause is too high? Can you perform NAME simulations in which you alter the tropopause height to test this? How does the stability in the stratosphere compare to that in the troposphere? How does the stratospheric stability in NAME compare to that measured by the radiosonde?
23. Page 11, line 283. I couldn't find the part of section 2.1 in which you show that the interference of ash on the retrieval is significantly reduced.
24. Page 11, lines 280-287. There doesn't appear to be any discussion of figure 3c. Does this mean that this figure is not necessary? If so, it should be removed.
25. Page 12, figure 3. It is difficult to see this figure (printed in black and white) due to the coloured background used. Is this necessary?
26. Page 12, figure 3. What does the dashed line represent? Shouldn't it be a box with a latitudinal extent from 48-52N?
27. Page 13, line 295. You conclude that the underestimation is due to an underestimation of the fraction of SO2 released into the stratosphere. Could it also be due to overdispersion of SO2 in the stratosphere?
28. Page 13, 14, 15 and 16. A description of FSS and SAL metrics is already in the published literature. Is it necessary to include this in the main body of the text particularly in this highly idealised form? Since SAL is an object orientated verification metric it would be more informative to show a snapshot of a SO2 cloud with identified objects rather than the idealised example described in the text.
29. Page 13, lines 346-350. For some reason the writing changes to use first person pronouns. I'm not sure what ACP policy is but this section seemed to be written in a different style to the rest of the paper.
30. Page 17, figure 6. Please could the background colour be removed as it's difficult to distinguish the SO2 cloud when printed in black and white.

31. Page 17, line 399. Since the StratProfile has been designed to agree better with the TROPOMI SO2 cloud it's hardly surprising that it does.
32. Page 17, line 404. How do you identify the part of the cloud in the NAME simulation what is within the troposphere and stratosphere using the differences in the simulations? Are you assuming no exchange of SO2 from the stratosphere to the troposphere?
33. Page 21, table 3. This table includes the same information as shown in figure 9 I think but expanded for more thresholds. Could this be moved to the supplementary material?
34. Page 22, figure 10. Why is the SO2 mass lost too fast in the VolRes1.5 simulation? Why is the grey arrow in panel (b) the same as in (a)? Shouldn't it have smaller vertical extent to indicate reduced amplitude spread? Why does the arrow in panel (d) and (d) extend beyond the point with the most negative structure value? Is it necessary to show the individual TROPOMI overpasses? Perhaps including the daily averaged squares only would be easier to follow the evolution in the SAL scores?
35. Page 22, line 471. How do you know that the high VCDs are related to small-scale eddies?
36. Page 23, line 479. The authors state that the S-values are 'relatively close to 0'. Relative compared to what?
37. Page 23, line 480. Does the Location value have units or is it non-dimensional?
38. Page 23, line 483. '4-5 days after the start of the eruption'. Please refer to the dates used in the figure if possible.
39. Page 23, line 490. What happens after 5 days?
40. Page 23, line 493. What is the StratProfile simulation event 'better' than?
41. Page 23, line 499. What is the motivation for reducing the K_meso value by 75%? Is there evidence in the literature that this is appropriate or are you simply using K_meso as a tuning parameter?
42. Page 23, line 504. Why is the S score independent of the diffusion parameter? Is this because the size of the plume becomes greater than the size of the mesoscale eddies that K_meso is representing so the synoptic scale uncertainty dominates?
43. Page 24, figure 11. Here the AAI is used to indicate high concentrations of ash, thereby affecting the TROPOMI SO2 retrievals. This interference is referred to at several points earlier in the paper but until this figure I don't think any evidence was shown to support the statement that ash was potentially contaminating the retrieval. Perhaps this evidence could be included earlier in the paper?
44. Page 24, line 518. Why does the VolRes1.5 simulation loose SO2 mass at a much faster rate than TROPOMI?
45. Page 28, line 641. Do you change K_meso everywhere or just in the stratosphere?
46. Page 28, line 644. Why do you think that a 'precise value for the diffusion parameters' exists? Turbulence is typically patchy, suggesting that a constant value is unsuitable.
47. Page 29, lines 681-684. This appears to be a repetition of the results already stated in the results section, not a conclusion.
48. Page 30, line 695. It is surprising that no reference to the extensive literature on stratospheric dispersion is included here.