1) The addition of your new figure 1 made the method more clear, yet the figure was not sufficiently described in the text. If I understand correctly, the thick, solid segments on the CALIOP orbital tracks (do not call them "swaths"!) are the segments where volcanic aerosol was detected, right? These are then selected for further analysis, i.e., as starting points for trajectory analysis. The light blue cloud is presumably the SO2 plume detected during successive AIRS swaths, and the grey bits the areas where FLEXPART trajectories lead to plume altitude information, correct? Please explain this in the text. Then link Figures 1 and 4 (Sect. 3) to make the method explanation more complete We have changed the figure text according to your suggestion. We have linked it to Section 3 and to Fig. 4, and refer to the concept sketch throughout the text.

2) "tens of degrees Celsius" -> "several tenths of a degree Celsius"
We changed accordingly.

3) Your method yields results that differ markedly from all other data in Fig. 8. Is there evidence that your method is better than the others? Please explain in more detail why the results are so different, as simply stating that the other results are based on different data sets is not sufficient. Why did Wu et al. not see the two peaks and why do they strongly underestimate the total amount of SO2? The Haywood and Mills profiles seem to be much closer to what you found than Höpfner or Ge. How does that influence model results? This is actually a very important point, because if you want modelers to use your computationally expensive method for investigations of the (radiative) effects of volcanic eruptions, you'll need to convince them that it is worth it. So: what do you think is the improvement that can be achieved by using your method or data?

We have rewritten two paragraphs in the Discussions section. We now discuss all datasets and included a more thorough discussion on our dataset. We also clarified that simulations with higher vertical resolution should result in more realistic model simulations.

Wu et al. had only slightly lower SO2 mass in the stratosphere than we have (0.9 Tg vs. 1.1 Tg). They don't see any big peaks since their methodology differs markedly from ours. They use a trajectory model to infer the altitude distribution by computing trajectories at all altitudes from 0 - 20 km to find out which trajectories coincide with the measured horizontal extension of an SO2 cloud. That could lead to false positives, i.e. that altitude ranges not affected by volcanism could be transported in the same way as the volcanic layer, thus causing a broadening. In contrast we used altitude distribution from a lidar with 60 m vertical resolution the altitude range studied, which is about the best than can be obtained. Another difference is that Wu et al used altitude as their vertical parameter, whereas we used potential temperature. The relation between altitude and potential temperature varies over time and space, and the fact that transport in the stratosphere (in the first approximation) occur along potential temperature surfaces could cause additional broadening of the Wu et al profile. Wu et al., finally, validated their profile with an aerosol signal from the MIPAS instrument having, for the present purpose, the poor vertical resolution of 3 - 4 km.

When these issues have been resolved to my satisfaction, I will gladly accept the manuscript for publication in AMT.