

## Detailed point-by-point response to all referee comments

In the revised version of the manuscript attached below we have highlighted changes in red color, changes related to the comments from reviewer #1 in green and changes related to comments from reviewer #2 in blue. Hereafter we first provide an overview about the general changes made to the manuscript before we address the comments from the referees.

### General changes

All general changes are highlighted in red in the revised manuscript

In general, based on the referee comments, we have improved and slightly restructured the text.

In agreement with the editor, we have removed the entire Ambae section since IASI was the only reference products to compare our SO<sub>2</sub> LH results with and the SO<sub>2</sub> plume was pretty weak – only a few pixels exceeded our threshold criterion of SO<sub>2</sub> VCD > 20DU. We also did not find any CALIPSO or MLS data that could be used for the verification. First of all because there was no ash present and the plume was probably too weak and too low to be detected by MLS. Unfortunately the difference in the overpass times of TROPOMI and IASI were too big as to provide a meaningful comparison. Even with a trajectory model (HySPLIT) we were not able to correct for the different overpass times. This might be related to our limited knowledge of using trajectory models. We have therefore decided to expand our section on Sinabung and Sierra Negra and in addition show results of the very recent Raikoke eruption where we have several MLS, CALIPSO and IASI datasets available for validation.

We have added a new section that describes how the FP\_ILM algorithm could be integrated into the operational TROPOMI SO<sub>2</sub> retrieval

### Changes made based on review of anonymous referee #1 (Received and published: 14 March 2019)

All changes related to this review are highlighted in green in the revised manuscript. The response to the referee is structured in the following sequence: (1) comment from referee, (2) author's response, (3) author's change in the manuscript

- 1.1 One main concern, for the specific algorithm, is that it uses wavelengths as short as 310 nm from TROPOMI. While these wavelengths provide higher sensitivity to SO<sub>2</sub> and SO<sub>2</sub> plume height, straylight at these wavelengths also can impact the retrievals. I wonder how the authors address this in their training?
- 1.2 It is correct that straylight can introduce spectral features that may lead to bias in the retrieved SO<sub>2</sub> LH, but straylight in Band 3 of TROPOMI (i.e. around 310nm) is very small:  
According to Kleipool et al (2018), see <https://www.atmos-meas-tech.net/11/6439/2018/amt-11-6439-2018.html> table 8 and figure 25, bottom left, the in-band straylight after correction is as low as 0,5% (and thus meets the instrument requirements). Also no evidence for out-of-band-straylight in the UVIS was found by the TROPOMI L1 team (pers. comm. Quintus Kleipool)
- 1.3 We have added a comment on instrumental straylight in Sect. 3 (“Dependencies”)



- 2.1 I'm not sure how the 2 km accuracy stated in the abstract is determined. I assume this is based on results in Figures 2-4? On the other hand, the temporal mismatch between TROPOMI and the primary validation dataset (IASI) makes it difficult to support this accuracy estimate. Also given the difference in TROPOMI and IASI overpass time, some sort of trajectory analysis may help to better link the two retrievals in the comparison.
  - 2.2 The 2km accuracy is determined by applying the FP\_ILM to an independent test dataset with known SO<sub>2</sub> LH. Due to missing independent plume height data (from ground or other satellites) at exactly the same time as the S5p overpass, no other way to determine the accuracy is possible. Unfortunately we have only limited experience with trajectory analysis models, but we have performed simple particle trajectory analysis using HYSPLIT in which we have started particles at the IASI LH points and followed them until the S5P overpass time. Based on this 'forecast' we will make a collocation IASI-TROPOMI based on closest distance, and generate statistics (correlation, mean, std, ..) in the updated paper. This however works only for extended plumes, i.e. that of Sierra Negra and Ambae. For Sinabung, we are too close to the (active) source to perform a meaningful collocation.
  - 2.3 We have added a comment on how we come to the conclusion of achieving the 2km at all occurrences in the text.
- 
- 3.1 For NRT applications, the algorithm would require input of O<sub>3</sub> VCD? How much time is required to retrieve O<sub>3</sub>? Also can the authors discuss if the plume height algorithm will run all the time or just be triggered by eruptions?
  - 3.2 Yes, the O<sub>3</sub> VCD is required to compensate for the strong spectral interference between SO<sub>2</sub> and O<sub>3</sub> in the spectral range used for the SO<sub>2</sub> LH retrieval. The O<sub>3</sub> VCD is an operational TROPOMI product and is also used as input for the operational SO<sub>2</sub> VCD retrieval (needed to calculate AMFs). Hence no additional time to retrieve O<sub>3</sub> would be required in the operational TROPOMI L2 retrieval. The SO<sub>2</sub> LH algorithm should be triggered by the operational SO<sub>2</sub> VCD detection flag, which flags pixel showing an enhanced SO<sub>2</sub> signal. This can of course be optimized to trigger the LH retrieval above a certain VCD threshold. We will add a paragraph, discussing the possible implementation in the operational TROPOMI retrieval.
  - 3.3 We have added a new section in which we discuss a possible implementation into the operational S5P environment, see Section 5 "Implementation in an operational environment"
- 
- 4.1 Page 1, Line 15: oxidation of SO<sub>2</sub> also takes place in the troposphere.
  - 4.2 Indeed. We have corrected the sentence slightly
  - 4.3 We have corrected the sentence in P1 L16
- 
- 5.1 Page 3: Line 30: can the authors give some examples of the parameter grids determined by the smart sampling technique?
  - 5.2 Copy paragraph from previous paper (?)
  - 5.3 We have added a reference to the previous paper on P4 L8, where this is explained in detail in order to avoid repeating the same examples. Also a reference to the original paper presenting this technique has been added
- 
- 6.1 Page 4, Line 12: Maybe figure 6 should be Figure 1, since it is discussed before all other figures.

6.2 You are right. We will show Figure 6 first before the other figures

6.3 Figure 6 has been moved to the front

7.1 Page 6, Line 10: is there a way to determine which rows should be used for training that would provide optimized retrievals? One would assume that pixels with large SO<sub>2</sub> VCDs should be used?

7.2 In principle, an inversion operator for each single row should be trained, since each row has its own instrumental characteristic. For the detection of a volcanic SO<sub>2</sub> plume, we don't know in advance in which detector row it will be detected, so we made the choice of interpolating the plume height results retrieved for every 50<sup>th</sup> row to the actual row where it was detected.

7.3 We have added a comment to P5 L19

8.1 Figure 1: it is a bit surprising that the error can be larger for really high SO<sub>2</sub> VCD (close to 1000 DU), do we know why?

8.2 Note that the figure shows the results for random parameter choices. The large errors seen for really high SO<sub>2</sub> VCD is related to a very high SZA in these cases

8.3 The figure has been updated (it is now Figure 2), now showing the dependencies for the first an optimized training test data

9.1 Figure 1: it is not very obvious from the figure that high albedo values have a negative impact on the height retrievals – why limit training to albedo < 0.5?

9.2 The large ranges of albedo introduces large variations in the spectra (many reflections from the surface). This additional variations correlate with the variations due to perturbations in the SO<sub>2</sub> VCD and LH. Therefore, to make the algorithm more stable, we restrict the albedo range to the physically relevant cases.

9.3 We added a sentence to clarify our approach in Section 3 (“Dependencies”)

10.1 Figures 3 and 4: add shade to mark +/- 2 km from the “real” plume height.

10.2 Very good suggestion!

10.3 We have updated the Figures, now showing the suggested shade to mark the 2km uncertainty

11.1 Figure 7: plume height retrievals were also done for pixels with small amount of SO<sub>2</sub>?

11.2 Yes, we have applied the FP\_ILM to all pixels for which an enhanced SO<sub>2</sub> signal was detected (there is a flag in the operational SO<sub>2</sub> product) are, hence even for low SO<sub>2</sub> VCDs with correspondingly lower SO<sub>2</sub> LH accuracies

11.3 We have updated all figures such that only pixels with VCD>20DU are shown in the LH plots

12.1 Figure 8: would suggest to only plot TROPOMI retrievals with SO<sub>2</sub> > 20 DU.

12.2 We also thought about this, but then one would only see a few pixels and not the entire extend of the SO<sub>2</sub> plume

12.3 We have removed the entire Ambae section. Nevertheless, for the final version we have decided to plot in all cases the entire plume for the VCD plot, but only the SO<sub>2</sub> LH for pixels with >20DU

- 13.1 Figure 9: how much can the ~30 km retrieved plume height be trusted, if the training data only go up to 20 km (Table 1)?
- 13.2 It is correct that the any LH result exceeding the training range has probably a high uncertainty. These cases should be treated with different optimized retrieval operator in an operational environment. In this sense we propose a 2-step retrieval: Once extreme values are detected, we switch to a designated operator.
- 13.3 We have removed the entire Ambae section, since we did not find a good match between Tropomi and IASI (too long time difference), which we were not able to correct for using HySplit updated the figures accordingly. Also other datasets (CALIPSO, MLS) did not provide any measurements for the Ambae eruptions. Instead, we added (in agreement with the editor) a new subsection on the Raikoke volcanic eruption, for which we have an extended set of independent sources of the LH
- 14.1 Figure 11: suggest to plot CALIPSO ground track on one of the maps.
- 14.2 Good idea, we will overplot the CALIPSO ground track
- 14.3 In all SO<sub>2</sub> LH plot we are now plotting the CALIPSO ground tracks, if applicable

## Changes made based on review of anonymous Referee #2 (Received and published: 15 May 2019)

All changes related to this review are highlighted in blue in the revised manuscript. The response to the referee is structured in the following sequence: (1) comment from referee, (2) author's response, (3) author's change in the manuscript

- 1.1 Overall I think the paper could be suitable for publication in AMT after some moderate revisions. The structure of the paper could be improved – currently there are many short paragraphs and not all the information is presented in a logical order, and many figures could be improved (see detailed comments below).
- 1.2 We will improve both text and figures
- 1.3 The entire text has been improved and partly restructured (see blue highlighted text) and the figures now show zoom-ins.
  
- 2.1 I do question why the authors only simulated SO<sub>2</sub> layer heights up to a maximum of 20 km? Major volcanic eruptions (with the largest potential climate impacts) can inject SO<sub>2</sub> to greater altitudes and hence it would be interesting to see how the FP\_ILM algorithm would perform in such a scenario, given that the FP\_ILM retrievals of SO<sub>2</sub> at 20 km altitude appear least accurate for low SO<sub>2</sub> VCDs (e.g., Figure 4 suggests that a VCD > 40 DU is needed for accurate retrieval).
- 2.2 We will extend our training dataset to include SO<sub>2</sub> LH up to 30km. We will also show how the NN responds to SO<sub>2</sub> LHs not used in the training dataset.
- 2.3 The training dataset has been extended to cover also SO<sub>2</sub>LH up to 30km. In order to optimize our algorithm, we however have restricted our training to use only LHs up to 25km (see Table 1). Although also higher LHs from strong volcanic eruptions occur sometimes, using a broad training data range also has an influence on the accuracy of the retrieval, since the algorithm has to cope with this broad range. To a limited extend however, the FP\_ILM is also able to extrapolate to an untrained parameter range, however with significantly decreasing accuracy. We have added a comment in Sect. 3 (“Dependencies”)
  
- 3.1 On a related note, under very high SO<sub>2</sub> loadings in a major eruption the ozone (O<sub>3</sub>) VCD retrievals may be inaccurate (due to SO<sub>2</sub> interference), and I assume this would preclude accurate SO<sub>2</sub> LH retrieval (since the O<sub>3</sub> VCD is a required input).
- 3.2 Small errors are expected in O<sub>3</sub> VCD and therefore the impact on the SO<sub>2</sub> LH retrieval should be minor. See Lerot et al (2013, “Homogenized total ozone data records from the European sensors GOME/ERS-2, SCIAMACHY/Envisat, and GOME-2/MetOp-A”): “The effect is negligible, except for a major volcanic eruption scenario with SO<sub>2</sub> column amounts exceeding 50 DU. In this case, total ozone errors may reach a few percent.”
- 3.3 We have added a comment on the dependency of our SO<sub>2</sub>LH retrieval on error in the O<sub>3</sub>VCD in Sect. 3 (“Dependencies”)
  
- 4.1 I also find that the stated SO<sub>2</sub> LH accuracy of ‘better than 2 km for SO<sub>2</sub> VCD > 20 DU’ is a little exaggerated, especially for higher SO<sub>2</sub> LH, e.g., for an SO<sub>2</sub> VCD of 20 DU at 20 km, the SO<sub>2</sub> LH appears underestimated by ~ 5 km in Figure 3.

- 4.2 We will update the figure, clearly showing the anticipated accuracy of 2km. Furthermore we will compare our results to IASI and MLS results by determining the LH at the same overpass time using an dispersion model (i.e. HySplit)
- 4.3 We have updated Fig. 3, now showing much more test data points for pre-defined LHs. This figure also proves that we reach the 2km uncertainty using this independent test dataset. We have decided to not use HySPLIT to compare TROPOMI and IASI data since our experience with trajectory models is very limited. We have therefore performed comparisons with MLS and CALIPSO data and only compare to IASI data on a qualitative level.
- 5.1 Another weakness is the validation of the TROPOMI SO<sub>2</sub> LH using IASI. Since IASI measurements are not coincident with TROPOMI, only broad conclusions can be drawn from the comparisons. In addition to CALIOP, the authors could explore the use of Microwave Limb Sounder (MLS) SO<sub>2</sub> data from the Aura satellite to validate the SO<sub>2</sub> LH retrievals.
- 5.2 Good point. We will check whether we find MLS measurements for the volcanic cases studied. Unfortunately, IASI is the only other satellite source which we can use for the validation. In order to correct for the overpass time difference between S5P & IASI, we will use a trajectory model (HySplit) to forecast the IASI LH at the S5P overpass time
- 5.3 We have added comparisons with MLS data and found additional CALIPSO overpasses that we could use for comparison. We decided to not use HySPLIT to correct for the overpass time difference between TROPOMI and IASI due to our limited experience with trajectory models. Comparisons with IASI data is used for a qualitative validation on the accuracy of retrieved SO<sub>2</sub> LHs from TROPOMI measurement
- 6.1. P1, L21: there are many different ‘flavours’ of DOAS algorithm, so writing ‘the DOAS algorithm’ seems to be a major generalization. Furthermore, there should probably also be a reference to the first multi-spectral Total Ozone Mapping Spectrometer (TOMS) SO<sub>2</sub> retrievals, which used a different approach.
- 6.2. We will update the text accordingly and add a reference to TOMS and other SO<sub>2</sub> retrievals (e.g. OMI PCA)
- 6.3. We have updated the text (see P2, L1) and now also list the PCA and Krueger-Kerr algorithms to retrieve SO<sub>2</sub>
- 7.1 P2, L1: ‘fast enough for NRT retrievals’ – algorithm speed/computational cost and the timeliness of retrievals are mentioned several times in the paper (e.g., P2, L19-20; P3, L3-5; P3, L17), but there is little quantitative information (I see that there is some information on P6). I would recommend adding a brief discussion to the introduction describing the data latency desired (e.g., for aviation safety and other applications) and estimates of current processing speeds.
- 7.2 Very good point. The TROPOMI/S5P NRT data is available 3 hours after sensing. We will add a sentence describing the current processing speed of the operational SO<sub>2</sub> VCD retrieval along with an estimation of the extra time needed for the SO<sub>2</sub> LH retrieval
- 7.3 We have added the new Section 5 “Implementation in an operational environment” that describes how the FP\_ILM could be integrated in the operational TROPOMI SO<sub>2</sub> retrieval and also provide detailed information about the speed of the algorithm with respect to current processing speeds.

- 8.1 P2, L2-4: I think any SO<sub>2</sub> algorithm (regardless of whether an AMF is explicitly used) needs to make some assumptions regarding the SO<sub>2</sub> vertical distribution (due to the pressure/ temperature dependence of SO<sub>2</sub> absorption).
- 8.2 Indeed, we will rephrase the sentence
- 8.3 We have rephrased the sentence, see P2 L3-5
  
- 9.1 P2, L5: accurate AMF calculations could also include parameters such as cloud fraction, surface pressure and surface reflectivity. Also, it is not unique to the 305-335 nm range.
- 9.2 Indeed, we will rephrase the sentence
- 9.3 We have rephrased the sentence, see P2 L6
  
- 10.1 P2, L6: some of paragraph 2 basically restates the previous paragraph; i.e., the SO<sub>2</sub> VCD is strongly dependent on the vertical distribution of SO<sub>2</sub>, as the latter strongly affects the AMF. These paragraphs could be reorganized/combined to clarify the text.
- 10.2 We will reorganize the text
- 10.3 We have restructured the text for clarification
  
- 11.1 P2, L10: it could be added here (instead of L16) that the usual approach for operational SO<sub>2</sub> retrievals (not only from TROPOMI but also other UV sensors) is to assume several different a-priori SO<sub>2</sub> vertical distributions and provide VCDs for each. I think it is also important to stress that above ~5 km or so (i.e., in the upper troposphere and above), the vertical SO<sub>2</sub> distribution has relatively little impact on the VCD (although the actual altitude is still of interest of course).
- 11.2 We will update the text accordingly
- 11.3 The text has been updated, see P2, L15-18
  
- 12.1 P2, L14: it is not only the number of photons but also the UV wavelengths interacting with the SO<sub>2</sub> layer that are influenced by the SO<sub>2</sub> layer height.
- 12.2 Correct. We will update the text accordingly
- 12.3 We have updated the text accordingly
  
- 13.1 P2, L16: TROPOMI is first mentioned here, but some key information is provided later on L29 – some reorganization is needed.
- 13.2 We will reorganize the paper to improve readability
- 13.3 We have removed the TROPOMI reference here (see comment 11 above)
  
- 14.1 P2, L21: Extensive -> Extended.
- 14.2 We will update the text accordingly
- 14.3 The text has been updated
  
- 15.1 P2, L22: it is not clear what is meant by 'strong volcanic eruptions'. Eruptions can be relatively weak and still produce high SO<sub>2</sub> column amounts, and vice versa.
- 15.2 We will use the term Volcanic Explosivity Index (VEI) instead to clarify what is meant here

- 15.3 We are now using the term VEI throughout the paper to characterize the strength of the volcanic eruptions
- 16.1 P2, L26-27: is there a reference to support the statement that IR SO<sub>2</sub> height retrievals are more accurate than UV retrievals? I'm not sure that either approach has been extensively validated.
- 16.2 The paper of Clarisse et al. (2014, see <https://www.atmos-chem-phys.net/14/3095/2014/>) on the Nabro eruption demonstrates that the IR is capable of being very accurate and sensitive to SO<sub>2</sub> height when the SO<sub>2</sub> columns are low. This paper also presents quite comprehensive validation with CALIPSO. Another (but similar) algorithm by the Carboni et al. (2016, see <https://www.atmos-chem-phys.net/16/4343/2016/>) shows equally good results. We will add these two references to our text
- 16.3 Both references have been added to the text (see P2, L30)
- 17.1 P3, L3: It would be useful to know how often the algorithm needs to be 'trained'. Is re-training necessary if the TROPOMI data quality changes, or for other reasons?
- 17.2 The algorithm needs to be re-trained only when large changes in the instrument slit function (ISRF) or SNR occur. Based on the experience from previous UV satellite sensors, large ISRF changes are not expected and, as can be seen from Figure 4, a (moderate) change in the SNR has no big impact on the SO<sub>2</sub> LH retrieval results.
- 17.3 We have added a comment on P5, L21
- 18.1 P3, L6: please also cite the first paper on the FP\_ILM SO<sub>2</sub> LH algorithm here. It might also be useful to briefly summarize the 'improvements' to the algorithm here too.
- 18.2 We will add the reference here
- 18.3 We have added the reference as well as a description about the improvements on P3, L22-26
- 19.1 P3, L24: 'plume profile'.
- 19.2 We will update the text
- 19.3 We have updated the text
- 20.1 P4, L10: eight parameters were used to simulate the spectra, but a larger number (10) of PCs is needed to retrieve the layer height. This seems to contradict the authors assertion (P4, L9) that 'fewer parameters' are used to characterize the dataset after the PCA. Some more explanation/clarification may be needed here.
- 20.2 The 'fewer parameters' refers to the comparison between 10 PCs and the corresponding 161 spectral points of the SO<sub>2</sub> LH fitting window.
- 20.3 We have clarified our approach in Sect. 2
- 21.1 P4, L12: Figure 6 is the first Figure referenced here – in which case the Figures should be reordered.
- 21.2 We will reorder the Figures accordingly.
- 21.3 Figure 6 has now moved to Figure 1



- 22.1 P4, L16: This paragraph (and also the following one) is probably difficult to follow for anyone not acquainted with neural networks or machine learning. Several new terms are introduced without elaboration (loss function, weight vectors, hidden layers). I recommend that the authors provide more details on the procedure.
- 22.2 We will update the paragraph and add more information about NNs and machine learning
- 22.3 We have updated the text and improve more detailed information about the NN approach in Sect. 2
- 23.1 P5, L5: shouldn't the O3 VCD also be listed as a direct dependency?
- 23.2 That's correct, we will add it to the list
- 23.3 O3 VCD has been added to the list
- 24.1 P6, L2: from Figure 4, it appears that the accuracy of SO<sub>2</sub> LH retrieval does not significantly improve with increasing SNR for high altitude SO<sub>2</sub> LH (20 km). Can the authors explain this?
- 24.2 The SO<sub>2</sub> LH retrieval is more sensitive to higher-altitude plumes than to low-altitude plumes. Therefore, increasing SNR will improve the accuracy for low plumes but not so much for higher plumes.
- 24.3 The SNR has only a minor impact on our SO<sub>2</sub>LH accuracy, since we effectively filter out noise using the PCA approach and using only the first 10 PC scores. We only see an influence on the results for very low SNR < 500. We have added a comment in Sect. 3, P6 24-27.
- 25.1 P6, L23: The IASI data are not the only source of independent SO<sub>2</sub> LH data. The Microwave Limb Sounder (MLS) on the NASA/Aura satellite can provide some information on SO<sub>2</sub> LH, albeit with limited spatial coverage and vertical resolution. The afternoon MLS overpass is nearly coincident with TROPOMI, and MLS did detect at least some of the eruptions discussed in the paper. I wonder if the authors considered using the MLS data to validate their SO<sub>2</sub> LH retrievals?
- 25.2 Good point. We will also include MLS results for Sinabung
- 25.3 We have added detailed MLS results for Sinabung, Sierra Negra and Raikoke
- 26.1 P6, L26: it would be useful to have at least a few more details on the Ambae volcanic activity (and actually for all the eruptions discussed in the paper), e.g., from the Smithsonian Institution Global Volcanism Program reports.
- 26.2 Ok, we will add more information on the volcanoes presented in this paper
- 26.3 As already described in the general comments above, we have removed the entire Ambae section. However, for all other volcanoes, we have added more information on the volcano itself
- 27.1 P6, L30: it is unclear why the plume is 'aged'? See general comment below regarding the SO<sub>2</sub> LH map (only pixels with robust LH retrievals should be shown). The two plumes discussed in the text are swamped by areas of blue (low SO<sub>2</sub> layer heights which I presume are incorrect due to the generally low SO<sub>2</sub> VCDs) and hence hard to see.
- 27.2 With 'aged' we mean that the plume has already travelled from its source after the eruption until the S5P measurement took place. Actually we discovered that the image shows the plume with a time difference of 24 hours due to the date line: The plume in Fig.9 close to the volcano (around 180E) is the plume observed in the morning of 27 June, whereas the part of the plume around 150W

is the same SO<sub>2</sub> cloud but observed on the evening of 27 June. We will hence update the figure accordingly. In the revised paper we will only show results for SO<sub>2</sub> close to 20DU.

27.3 The Ambae section has been removed. We now only show SO<sub>2</sub>LH maps for SO<sub>2</sub>>20DU.

28.1 P7, L6: need to stress that this is also from Ambae.

28.2 This sentence belongs to the Ambae section, so this should be clear. But we will make this more clear in the text

28.3 The Ambae section has been removed

29.1 P7, L8: acid rain is not usually an issue for stratospheric SO<sub>2</sub>.

29.2 Indeed. We will correct the sentence

29.3 The Ambae section has been removed

30.1 P7, L15: a brief description and reference for the CALIOP instrument is needed (also provide the full name of the sensor).

30.2 We will add a brief description for CALIOP

30.3 Details on CALIOP have been added to Sect. 4 (see P7, L26)

31.1 P7, L16: it should be noted that it is not necessarily the case that SO<sub>2</sub> and ash/aerosols are collocated, as gas and ash can separate in volcanic clouds (e.g., as ash falls out to lower altitudes).

31.2 Correct. We will point this out in the revised paper

31.3 We have added a comment on P7 L27

32.1 Figures 3 and 4 are quite similar and could perhaps be combined as one figure.

32.2 That's correct. We will combine the figures

32.3 We have updated Figure 3 & 4 showing more datapoints from the test dataset. A combination of the figures would result in too many datapoints

33.1 Figure 5: this figure is a bit cluttered and could perhaps be improved by removing some of the data from the plot, e.g., using just the higher SO<sub>2</sub> VCDs.

33.2 We will improve the figure

33.3 We have improved the figure by using smaller symbols

34.1 Figures 7-15: General comment on the SO<sub>2</sub> map figures: I would recommend 'zooming in' as much as possible on the SO<sub>2</sub> plumes to show the detailed structure (especially Sinabung). Also, I think the SO<sub>2</sub> LH plots should only show those TROPOMI pixels with robust SO<sub>2</sub> LH retrievals (i.e., SO<sub>2</sub> VCD > 10-20 DU or so), since otherwise most of the plots are showing invalid data.

34.2 We will make zoom-ins. The original intend was to show the same Lat-Lon range as when overplotting the IASI data.

34.3 We have updated all SO<sub>2</sub> maps and zoom-in to the volcanic plume(s)

- 35.1 Figure 8: I'm not sure that it is necessary to show both the IASI-A and IASI-B SO<sub>2</sub> LH data. Since neither are coincident with TROPOMI, just show the overpass that is closest in time and/or which has the best coverage of the volcanic plume.
- 35.2 Good point. Another idea would be to use the IASI data as input to a transport model (e.g. HySplit) and forecast the plume movement for the S5P overpass time to perform the comparison
- 35.3 We have removed the Ambae section
- 36.1 Figure 11: the CALIPSO satellite track corresponding to the lidar data in Figure 12 should be shown on the maps.
- 36.2 We will add the ground track to the figure
- 36.3 We have added CALIPSO ground tracks to all SO<sub>2</sub>LH figures when we compare with the CALIPSO data
- 37.1 Figure 12: this figure is also not very clear. I recommend 'zooming in' to the volcanic plume to show the data more clearly, and only plotting the red symbols (SO<sub>2</sub> VCD > 20 DU). It is also not clear which features represent the Sinabung volcanic eruption cloud and which are meteorological clouds; this could be highlighted on the plot.
- 37.2 We will update the plot and zoom-in to the volcanic plume
- 37.3 We are now only showing SO<sub>2</sub> LH for SO<sub>2</sub> > 20DU. As we are not that familiar with CALIPSO L1 data plotting we prefer to stick to the official CALIPSO images from NASA and simply overplot our data without zooming in. We have also decided to show the Vertical Feature Mask instead of the attenuation signal, making it easier to identify different features
- 38.1 Figure 13: the left-hand panel does not seem to show much if any useful IASI data. Given that there are CALIOP data for this case, perhaps the IASI data are not needed and this figure could be removed.
- 38.2 Good point, we will remove the left panel
- 38.3 We have decided to not show IASI data for the Sinabung eruption

# SO<sub>2</sub> Layer Height retrieval from Sentinel-5 Precursor/TROPOMI using FP\_ILM

Pascal Hedelt<sup>1</sup>, Dmitry S. Efremenko<sup>1</sup>, Diego G. Loyola<sup>1</sup>, Robert Spurr<sup>2</sup>, and Lieven Clarisse<sup>3</sup>

<sup>1</sup>Remote Sensing Technology institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany

<sup>2</sup>RT Solutions Inc., Cambridge, MA, USA

<sup>3</sup>Université libre de Bruxelles (ULB), Service de Chimie Quantique et Photophysique, Atmospheric Spectroscopy, Brussels, Belgium

**Correspondence:** P. Hedelt (pascal.hedelt@dlr.de)

**Abstract.** Accurate determination of the location, height and loading of SO<sub>2</sub> plumes emitted by volcanic eruptions is essential for aviation safety. The SO<sub>2</sub> layer height is furthermore one of the most critical parameters that determine the impact on the climate. Retrievals of SO<sub>2</sub> plume height have been carried out using satellite UV backscatter measurements, but until now, such algorithms are very time-consuming. We have developed an extremely fast yet accurate SO<sub>2</sub> layer height retrieval using the Full-Physics Inverse Learning Machine (FP\_ILM) algorithm. This is the first time the algorithm has been applied to measurements from the TROPOMI instrument on board the Sentinel-5 Precursor platform. In this paper, we demonstrate the ability of the FP\_ILM algorithm to retrieve SO<sub>2</sub> plume layer heights in near-real-time applications with an accuracy of better than 2 km for SO<sub>2</sub> total columns larger than 20 DU. We present SO<sub>2</sub> layer height results for the volcanic eruptions of Sinabung in February 2018, Sierra Negra in June 2018 and Raikoke in June 2019, observed by TROPOMI.

10 *Copyright statement.* TEXT

## 1 Introduction

Global satellite observations allow for the timely detection and monitoring of SO<sub>2</sub> emitted from volcanic eruptions, even in remote regions, where no ground-based instruments are installed (see e.g. Fioletov et al. 2013). Satellite measurements of UV earthshine spectra in the wavelength range between 305 and 335 nm provide the highest sensitivity to SO<sub>2</sub> in the Earth's atmosphere. Volcanic eruptions can inject large amounts of sulphur dioxide (SO<sub>2</sub>) into the atmosphere, where it is either subject to dry and wet deposition, as well as oxidization within a few days in the troposphere (see e.g. Lee et al. 2011 or Myles et al. 2011), or oxidization over a period of several weeks to sulfate aerosols in the stratosphere (see e.g. Robock, 2000; Forster et al., 2007 and von Glasow et al. 2009). Sulfate aerosols can affect the Earth's radiative forcing and have an impact on clouds (see e.g. McCormick et al., 1995; Robock, 2000 and Malavelle et al. 2017).

Based on UV earthshine measurements, SO<sub>2</sub> vertical column densities (VCDs) can be retrieved easily using for example the Differential Optical Absorption Spectroscopy (DOAS) technique, see e.g. Rix et al. (2012), a Principal Component Analysis

(PCA), see e.g. Li et al. (2017), or the Krueger-Kerr algorithm, which was applied to retrieve SO<sub>2</sub> from NASA TOMS, see e.g. Krueger et al. (1995). These methods are fast enough for near-real time (NRT) retrievals. Nevertheless, all algorithms retrieve only the slant column - in order to calculate the VCD a conversion factor (called the airmass factor, AMF) has to be applied, which includes explicit or implicit assumptions about the vertical distribution of SO<sub>2</sub> in order to determine the effective light path. Note that AMFs are calculated by means of multiple-scattering radiative transfer models assuming known vertical distributions of SO<sub>2</sub> and O<sub>3</sub>, as well as viewing, surface and cloud properties.

Unfortunately, the vertical distribution (in terms of the plume layer height) of SO<sub>2</sub> is usually unknown at the time of the measurement since it is not easy to extract from the spectral signature (see Yang et al. 2009 and Nowlan et al. 2011): The SO<sub>2</sub> loading (VCD) has a direct effect on the optical depth, whereas the layer height (LH) has an indirect effect on the optical depth since it influences the number of photons passing through the SO<sub>2</sub> layer, the UV wavelengths interacting with the SO<sub>2</sub> layer, as well as the layer optical depth due to the temperature dependency of the SO<sub>2</sub> absorption crosssections (Yang et al., 2009).

Even if there are ground-based or aircraft measurements of the SO<sub>2</sub> layer-height (LH), the data are generally difficult to use for validation, since e.g. for strong eruptions (Volcanic Explosivity Index VEI $\geq$ 3), volcanic plumes are typically transported over long distances and the number of collocations is small. Thus, for volcanic SO<sub>2</sub> measurements, the vertical distribution of SO<sub>2</sub> is a key parameter limiting the product accuracy. The usual approach for operational SO<sub>2</sub> retrievals is to assume several different a-priori SO<sub>2</sub> vertical distributions and provide VCDs for each (see e.g. Theys et al. 2017) along with an averaging kernel (AK), in order that the user can calculate the VCD for an arbitrary SO<sub>2</sub> vertical distribution. However, in the upper troposphere and above the vertical SO<sub>2</sub> distribution has little impact on the VCD.

To date, SO<sub>2</sub> layer-height retrievals have used computationally demanding direct fitting inversion methods, which are not suitable for NRT applications. For the retrievals based on satellite UV measurements, Yang et al. (2009) and Yang et al. (2010) developed an Extended Iterative Spectral Fitting (EISF) algorithm for the Ozone Monitoring Instrument (OMI) aboard the NASA Aura satellite, and Nowlan et al. (2011) introduced an optimal estimation (OE) scheme for the GOME-2 instrument (Global Ozone Monitoring Experiment-2) aboard the EUMETSAT/ESA MetOp satellite fleet. For strong volcanic eruptions (VEI $\geq$ 3), the accuracy of the retrieved SO<sub>2</sub> LH using this approach is in the range 0.5–1 km, whereas for small SO<sub>2</sub> absorption it is around 2 km (see Nowlan et al. 2011).

Satellite infrared sounders also offer the opportunity to measure both SO<sub>2</sub> VCDs and LH (see Clarisse et al., 2008; Carboni et al., 2012; Clarisse et al., 2014 and Carboni et al. 2016) using optimal estimation algorithms for the Infrared Atmospheric Sounding Interferometer (IASI) aboard the EUMETSAT/ESA MetOp satellite fleet. While infrared sounders have a inferior sensitivity to the lower tropospheric SO<sub>2</sub> than that for UV instruments, the layer height retrievals tend to have a better accuracy, and perform well even for low column amounts (up to the single-DU level, see e.g. Clarisse et al. 2014 or Carboni et al. 2016.), however, the horizontal resolution (12 km), is rather coarse.

In contrast, the TROPOMI instrument on board the Sentinel-5 Precursor satellite (S5P) launched on 13 October 2017 has a much higher spatial resolution of  $7 \times 3.5 \text{ km}^2$ , and will operate at an even smaller resolution of  $5.5 \times 3.5 \text{ km}^2$  beginning mid-August 2019. This allows us to observe and study SO<sub>2</sub> plumes at an unprecedented level of detail. Data turnover from

TROPOMI is very large, and this consideration will require the development of new retrieval schemes for the fast and accurate retrieval of SO<sub>2</sub> layer heights in an operational environment.

To this end, we have developed an algorithm called 'Full-Physics Inverse Learning Machine' (hereafter referred to as FP\_ILM) for the retrieval of the SO<sub>2</sub> LH based on satellite UV earthshine spectra. The FP\_ILM algorithm has been used for the retrieval of ozone profile shapes (Xu et al., 2017), the retrieval of surface properties accounting for **bidirectional reflectance distribution function (BRDF)** effects (Loyola et al., 2019), and the retrieval of SO<sub>2</sub> LH from GOME-2 (Efremenko et al., 2017). The algorithm creates a mapping between the spectral radiance and SO<sub>2</sub> LH using machine learning methods. The time-consuming training phase of the algorithm using radiative transfer model calculations is performed off-line, and only the inversion operator has to be applied to satellite measurements - this makes the algorithm extremely fast and it can thus be used in near-real time processing environments. In this second paper on the FP\_ILM SO<sub>2</sub> LH, we describe some improvements to the **original algorithm from Efremenko et al. (2017)**, and we apply it to a number of volcanic eruptions observed by TROPOMI during the operational phase **of the mission** (started April 2018).

The paper is organized as follows: In Section 2 we describe the improved FP\_ILM SO<sub>2</sub> LH algorithm. The sensitivity of retrieved SO<sub>2</sub> LHs to a number of different parameters is discussed in Section 3. In Section 4, the FP\_ILM is applied to S5P data to retrieve SO<sub>2</sub> LHs for selected volcanic eruptions. **Section 5 describes how the algorithm could be integrated in the operational TROPOMI SO<sub>2</sub> VCD retrieval algorithm.** We summarize the paper in Section 6.

## 2 FP\_ILM algorithm

Conceptually, the FP\_ILM consists of a training phase, in which the inversion operator is obtained using synthetic data generated with an appropriate radiative transfer (RT) model, and an operational phase, in which the inversion operator is applied to real satellite measurements. The main advantage of the FP\_ILM over classical direct fitting approaches is that the time-consuming training phase involving complex RT modeling is performed offline; the inverse operator itself is robust and computationally simple and therefore extremely fast. **In our previous paper (see Efremenko et al. 2017), we first introduced the FP\_ILM algorithm and applied it to GOME-2 observations of a number of volcanic eruptions. We used a combination of principal components analysis (PCA) and principal component regression (PCR) methods to train the inversion operator in order to retrieve the SO<sub>2</sub> LH. For the current paper we have improved the FP\_ILM algorithm with the use of a neural network (NN) approach, as outlined below.**

During the training phase, the Linearized Discrete Ordinate Radiative Transfer model (LIDORT) with inelastic rotational Raman scattering (RRS) implementation (Spurr et al., 2008) is deployed to compute simulated reflectance spectra in the wavelength range 310 – 335 nm. These spectra depend upon the following  $n = 8$  input parameters: the SO<sub>2</sub> VCD and LH, the surface albedo, the surface height, the O<sub>3</sub> VCD, the solar zenith angle (SZA), the viewing zenith angle (VZA) and the relative azimuth angle (RAA). **Table 1 provides a summary of the final parameter space after optimization (see below) used for the training of the final retrieval operator. Note that O<sub>3</sub> has to be included due to the strong spectral interference between SO<sub>2</sub> and O<sub>3</sub> in the spectral range considered. O<sub>3</sub> profiles are classified according to the total column amount of ozone, and the**

month and latitude zones as specified in the TOMS Version 8 O<sub>3</sub> profile climatology (Bhartia, 2003). The SO<sub>2</sub> plume profile is taken to have a Gaussian shape, characterized by the total SO<sub>2</sub> VCD loading and centered at a peak-concentration layer height  $z_p$ , along with a half-width fixed to 2.5 km. In the following, the retrieval of SO<sub>2</sub> layer height refers to the retrieval of the peak-concentration height  $z_p$ .

- 5 Simulations were done on a pressure/temperature/height grid from the US standard atmosphere, with a finer-grid vertical height resolution of 0.25 km below 15 km in order to resolve properly the Gaussian SO<sub>2</sub> plume shape. In total, some 131,072 simulated reflectance spectra have been calculated on a selective parameter grid established by means of a smart sampling technique developed by Loyola et al. (2016). Further details on the smart sampling technique applied to the SO<sub>2</sub> LH retrieval can be found in Efremenko et al. (2017).
- 10 The use of the LIDORT-RRS RT model is necessary, as it enables us to account for the effect of Raman scattering in the atmosphere: Solar irradiances exhibit strong Fraunhofer structures in this part of the UV spectral range, and earthshine spectra are characterized by the "filling-in" of Fraunhofer-solar and telluric-absorber features due to "inelastic" (wavelength-redistributed) rotational-Raman scattering by air molecules. For further details, see Efremenko et al. (2017) and Spurr et al. (2008).
- 15 The simulated high-resolution reflectance spectra are convolved with the TROPOMI Instrument Spectral Response Function (ISRF) v3.0.0 (released 2018-04-01)<sup>1</sup>. Note that the TROPOMI instrument comprises 450 rows, which are in principle single detectors with their own ISRFs. The signal-to-noise ratio (SNR) is about 1000 in our UV wavelength range. Thus, to account for instrumental noise in the training phase, uncorrelated Gaussian noise with a fixed SNR of 1000 is added to the simulated spectral data.
- 20 To extract the information about the layer height and to reduce the dimensionality of the spectral dataset, a PCA is applied to the simulated spectra. By thus characterizing the set of simulated measurements with fewer parameters, a simpler, more stable and computationally efficient inversion scheme can be realized. It was found that using 10 principal components (PCs) is sufficient to retrieve information about the SO<sub>2</sub> LH. These 10 PCs account for 99.994% of the spectral variance. The inclusion of additional PCs beyond 10 did not result in any improvements to LH retrievals, since higher-order PCs are increasingly
- 25 affected by noise. Figure 1 shows the ratio between the variance of the PCA-derived data set and the total variance of the complete spectra data set (the "explained" variance ratio) as a function of number of PCs included in the PCA.

The 10 PCs, together with the information about the O<sub>3</sub> VCD, the SZA/VZA/RAA angles, the surface pressure and albedo of each training data-point, are then used as input to train a feed-forward artificial neural network (NN) including regression (in our case a MultiLayer Perceptron Regression - MLPR), with the corresponding SO<sub>2</sub> LH as the output layer. Note here that

30 the SO<sub>2</sub> VCD is not part of the training, since it depends directly on the SO<sub>2</sub> layer height.

In general, NNs can be used for establishing a non-linear mapping between a dataset of numeric inputs and a set of numeric outputs. A NN consists of interconnected neurons (or nodes) that implement a simple, non-linear function (a sigmoid function in our case) of the inputs. Neural networking is a powerful tool for determining non-linear dependencies between datasets in

---

<sup>1</sup> Available here:

<http://www.tropomi.eu/data-products/isrf-dataset/>



remote sensing as shown by Loyola (2006). It consists of an input layer (representing the above mentioned input parameters), at least one so-called hidden layer and an output layer (representing the expected output - in our case the SO<sub>2</sub> LH). Each neuron in the hidden layer(s) transforms the values from the previous layer(s) with a weighted linear summation, followed by a non-linear activation function, which is in our case the sigmoid function. The output layer receives the values from the last hidden layer and transforms them into output values.

The training of the MLPR is an iterative process that tries to minimize the so-called loss-function (also known as the cost-function), which is a measure of how well a model predicts the expected outcome; at each time step, the partial derivatives of the loss-function with respect to the model parameters are computed in order to update the parameters. We note here that MLPR uses the "mean square error" loss-function. A regularization term is added to the loss-function that shrinks model parameters to prevent over-fitting: By building a complex neural network, it is quite easy to perfectly fit the training dataset. When this model is however evaluated on new data (here the satellite measurements), it performs very poorly. The regularization thus modifies the loss function by adding additional terms that penalize large weight vectors and preferring diffuse weight vectors.

After carrying out a PCA and MLPR parameter optimization using closed-loop retrievals to minimize differences between the retrieved and simulated layer heights, the final configuration for the neural network settled on the use of two hidden layers, with 32 nodes in the first layer and 10 nodes in the second.

In the operational phase, the first step is to use the principal component scores acquired during the training phase to transform a given TROPOMI spectral measurement data set to one with a lower dimension. Once this is done, the neural network inverse function is then applied to retrieve the SO<sub>2</sub> LH. In order to avoid the training of the NN for each of the 450 TROPOMI detector rows (with their own ISRF), we have trained the network only for every 50th detector row and interpolate the retrieved SO<sub>2</sub> LH results to the actual row where it was detected.

We note that the FP\_ILM algorithm only needs to be re-trained when large changes in the ISRF or SNR occur, see Section 3.

### 3 Dependencies

In this section, we study the dependency of the layer height retrieval on different parameters. We discriminate between direct dependencies (i.e. those parameters affecting the reflectance spectra) and indirect dependencies (i.e. affecting the training data and inversion algorithm) of the retrieved layer height.

- Direct dependencies: Viewing geometry, surface properties, ISRF, noise level (SNR), instrumental stray light, O<sub>3</sub> VCD
- Indirect dependencies: Number of layers in the neural network, number of PCs, parameter ranges for training

First, we train the FP\_ILM operator using 90% of the training dataset (see Table 1), and then apply the trained operator to the remaining 10% training spectra, for which we know the exact SO<sub>2</sub> LH. Figure 2 (red dots) shows the SO<sub>2</sub> LH difference as a function of solar/viewing geometry, SO<sub>2</sub> VCD, O<sub>3</sub> VCD, albedo and surface pressure, respectively. The figure shows clearly

a number of marked dependencies in the retrieved layer height, with notably high differences with respect to the real SO<sub>2</sub> LH for low and high layer heights, for low SO<sub>2</sub> VCD well as for high SZA.

Regarding the SZA dependency, a cutoff limit of 75° is usually set in operational SO<sub>2</sub> retrievals, because at high SZA the light path becomes very long and the noise level increases. Accordingly, we limit the SZA of the spectra used in the training to

5 SZA ≤ 75° in the following.

Clearly, for small SO<sub>2</sub> VCDs, the information content on LH in the spectral signature is very low. It follows that the inclusion of spectra with low SO<sub>2</sub> content in the training will have a negative effect on the entire neural network. In principle for lower SO<sub>2</sub> VCD loadings, more PCs could be included, but at some point the noise level signature will exceed that of the actual SO<sub>2</sub> absorption.

10 We have performed several tests in which we limit the training dataset by varying the allowed input parameter ranges. We have found an optimal parameter range that allows the retrieval of a broad range of SO<sub>2</sub> LHs even for low SO<sub>2</sub> VCDs. In the training phase, we use only spectra with SO<sub>2</sub> VCD ≥ 20 DU, surface albedo < 0.5 and SZA ≤ 75° to train the final inversion operator for TROPOMI. The albedo limit is set to 0.5, since large values of the albedo will induce large variations in the spectra (multiple reflections from the surface). These additional variations correlate with the variations due to perturbations in the SO<sub>2</sub> VCD and LH. Therefore, to make the algorithm more stable, we restrict the albedo range to the physically relevant cases. We have further limited the training values of SO<sub>2</sub> LH to the range 0–25 km. Although higher LHs from strong volcanic eruptions can occur, the use of a broad training data range also has an influence on the accuracy of the retrieval. To a limited extend however, the FP\_ILM is also able to extrapolate to an untrained parameter range, however with significantly less accuracy.

Using the optimized training dataset (see Table 1 for a summary of the parameter space), Fig. 2 (blue triangles) shows that the error on the retrieved SO<sub>2</sub> LH is less than 2 km and the dependency on SO<sub>2</sub> VCD is reduced. Figure 3 shows the retrieved layer height as a function of SO<sub>2</sub> VCD. As mentioned above, for low SO<sub>2</sub> VCDs, high-altitude layer heights cannot be retrieved - there is always a bias towards low layer heights. Only for SO<sub>2</sub> loadings in excess of 20 DU, do we retrieve layer heights with an uncertainty of less than 2 km.

To investigate the dependency of the retrieved SO<sub>2</sub> LH as a function of SNR, we have used two different noise levels (SNR of 500 and 1500). Figure 4 clearly shows that the SNR only has a minor effect on the accuracy of the retrieved layer height, with only slightly higher accuracy for increased SNR. This is to be expected, since we have used the first 10 PC scores and thus basically removed all noise features.

Concerning the dependency on the ISRF, we note that in the operational retrieval of SO<sub>2</sub> LHs from TROPOMI data, each detector row is effectively a single instrument with its own ISRF. The accuracy of the retrieved SO<sub>2</sub> LH can vary across detector rows. Figure 5 shows the SO<sub>2</sub> LH results when applied to the Sinabung eruption (see Section 4.1). Black dots show the LH for each 50th row, whereas the red cross shows the interpolated LH for the measurement row. Clearly, the retrieved LH is slightly different in each row (within 1 km) due to the different ISRF used for training the NN. Note that we determine the layer height for a set of fixed detector rows (for which we have trained the FP\_ILM separately), and then interpolate the layer height to the actual row. In this way we avoid jumps in the retrieved layer height between adjacent detector rows.

Instrumental straylight can introduce spectral features that may lead to bias in the retrieved SO<sub>2</sub> LH. However, according to Kleipool et al. (2018) the in-band straylight of TROPOMI after correction is as low as 0.5% and thus can be neglected. Furthermore no evidence for out-of-band-straylight was found (pers. comm. Quintus Kleipool).

During major volcanic eruptions ( $VEI \geq 3$ ) with very high SO<sub>2</sub> loadings, the O<sub>3</sub> VCD retrievals may be inaccurate due to SO<sub>2</sub> interference. However, this effect is negligible for weak eruptions. For strong eruptions (exceeding about 50 DU of SO<sub>2</sub>) the error on the O<sub>3</sub> VCD may reach a few percent (see Lerot et al. 2014) and hence the error on the SO<sub>2</sub> LH is negligible.

## 4 Application to TROPOMI data

Reflectance spectra from TROPOMI are determined from the operational L1 solar irradiance and earthshine radiance data (solar irradiance is measured on a daily basis). To correct for Doppler shifts between earthshine and irradiance spectra, we apply the wavelength calibration information from the operational L2 SO<sub>2</sub> product to first register the solar spectrum and then we use the fitted shift and squeeze parameters from the DOAS retrieval to calibrate the earthshine spectra. From this calibrated L1 data, we then calculate reflectances in the wavelength range 310 – 335 nm.

In the following subsections, we have applied the FP\_ILM operator to three major volcanic eruptions measured by TROPOMI. We have chosen eruptions having a peak SO<sub>2</sub> VCD exceeding our 20 DU threshold criterion and with an extended SO<sub>2</sub> plume that allows us to compare the FP\_ILM results with independent retrievals from other satellite data sources. In addition to L1 reflectances, we use information on SO<sub>2</sub> VCD, O<sub>3</sub> VCD, surface and viewing conditions from the operational SO<sub>2</sub> L2 product, see Pedernana et al. (2018).

For validation of our results, we have performed comparisons with independent MetOp/IASI SO<sub>2</sub> LH data (see Clarisse et al. 2014 for details), as well as SO<sub>2</sub> profile data from the Microwave Limb Sounder (MLS) on the NASA/Aura satellite (see Pumphrey et al. 2015). The MetOp platforms have different orbits from that of S5P, with widely different overpass times, so that direct satellite comparisons with IASI data will give only a qualitative validation on the accuracy of retrieved SO<sub>2</sub> LHs from TROPOMI measurements. The afternoon Aura/MLS overpass is nearly coincident with TROPOMI. MLS can provide some information on SO<sub>2</sub> LH, albeit with limited spatial coverage and vertical resolution, since it is only sensitive above about 147 hPa altitude (i.e. above around 13 km). We also checked on the availability of an overpass of the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO, Winker et al. 2010) satellite with the the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument. Note that CALIPSO measures only ash and aerosol absorption profiles. In this regard, we note that SO<sub>2</sub> plumes and ash or aerosols are not necessarily collocated, as gas and ash can separate in volcanic clouds. At the time of writing, IASI and MLS data are unfortunately the only sources for independent SO<sub>2</sub> LH satellite validation.

### 4.1 Sinabung

The Sinabung stratovolcano (2,460 m summit elevation) on the island of Sumatra has been highly active until September 2010 and was quiet until a new eruptive phase began in September 2013 which lasted until March 2018, see Venzke (2018b). On 19

February 2018, Sinabung erupted violently at around 02:55h UTC with its largest explosion to date, emitting a volcanic ash plume that rose to at least 16.8 km altitude and an SO<sub>2</sub> plume of up to 50 DU of SO<sub>2</sub> that was observed by several satellite instruments, including TROPOMI (see Fig. 6, overpass time around 06:30h UTC), OMI, IASI (overpass time around 03:30h UTC and 15:00h UTC), MLS (overpass time around 07:10h UTC). There was also an overpass of CALIPSO/CALIOP at about 07:15h UTC, over the volcanic plume.

The FP\_ILM algorithm retrieved SO<sub>2</sub> LHs extending up to 17 km, see Fig. 6; these show excellent agreement with the MLS SO<sub>2</sub> profile measurements inside the plume (see colored circles in Fig. 6), with peaks at 16.75 km close to the volcano, and 14.31 km in the northern part of the plume, see Fig. 8. Furthermore, our results are in close agreement with CALIPSO measurements, clearly showing an attenuation by ash or aerosols at altitudes around 15-18 km (color-coded in orange). Fresh volcanic plumes are typically rich in water vapour (especially for tropical eruptions), and thus the volcanic clouds also contain high concentrations of water droplets. Therefore, the classification in the CALIPSO vertical feature mask sometimes fails to pick up the volcanic ash or sulfate aerosol because of competing clouds. Nevertheless the altitudes of the identified features are likely those of the volcanic plumes themselves, given the collocations in time and space. We note that the brightness temperature difference plots (not shown) can help in the identification of an ash layer, since the difference in the brightness temperature between the CALIPSO channels at 10.6 μm and 12.05 μm becomes negative for ash, whereas for normal clouds it is positive.

In the figure we have overplotted our FP\_ILM SO<sub>2</sub> LH results in red dots. SO<sub>2</sub> LH retrieved for IASI/MetOp-A and -B measurements (not shown) indicated LH values at about 13 km for the 03:30h UTC overpass and LH up to 18 km around the overpass time of 15:00h UTC; these results agree well with the TROPOMI results.

Also the Pusat Vulkanologi dan Mitigasi Bencana Geologi<sup>2</sup> (PVMBG, also known as CVGHM) reported at 08:53h UTC a dark gray plume with a high volume of ash that rose at least to 16.8 km. Furthermore the Darwin Volcanic Ash Advisory Center<sup>3</sup> (VAAC) reported that LH values for these Sinabung ash plumes identified in satellite images, recorded by webcams, and reported by PVMBG continued to rise throughout the day to 13.7 km.

## 4.2 Sierra Negra

On 26 June 2018 a strong eruption at the Sierra Negra shield volcano (1,124 m summit elevation) located on Isabela Island (Galapagos) occurred. According to Venzke (2018a), this volcano has erupted several times since 1948, with the last eruption reported in 2005. After an increase in seismicity, Sierra Negra erupted at 20:09h UTC, producing a dense ash and SO<sub>2</sub> plume. The eruption was divided into an initial very energetic phase (VEI = 3) characterized by the opening of five fissures that lasted one day, and a long phase with lava flows from 27 July to 23 August 2018 with decreased gas emissions. TROPOMI was able to measure a very strong SO<sub>2</sub> plume (with loading in excess of 500 DU) only a few minutes (overpass at 20:12h UTC) after the start of the first eruption, as well as an extended SO<sub>2</sub> plume for several days after the eruption.

---

<sup>2</sup>see <http://vsi.esdm.go.id/>

<sup>3</sup>see <http://www.bom.gov.au/aviation/volcanic-ash/>

Figure 9 shows the SO<sub>2</sub> plume from 26 until 28 June at overpass times around 20:12h UTC. The corresponding FP\_ILM SO<sub>2</sub> LH is shown in Fig. 10. On 26th June, shortly after the eruption began, the retrieved SO<sub>2</sub> LH was around 4–6 km. On the following days however, a much higher SO<sub>2</sub> LH of around 14 km was retrieved by FP\_ILM, with some parts of the plume reaching about 18 km altitude. We note here that these high SO<sub>2</sub> LH values are not visible in the figure since they were retrieved for SO<sub>2</sub> VCD of about 10 DU and hence with low accuracy - in the figure only SO<sub>2</sub> LH results for SO<sub>2</sub> VCD ≥ 20 DU are shown.

We found only one closely related MLS SO<sub>2</sub> profile measurement intersecting the SO<sub>2</sub> plume on 27th of June, at a measurement time of 20:16h UTC, i.e. shortly after the TROPOMI measurement (see colored circle in Fig. 10 bottom, center). The MLS data in Figure 11 shows an SO<sub>2</sub> layer in the altitude range 9–12 km, which is in excellent agreement with the FP\_ILM SO<sub>2</sub> LH results of around 11–13 km at the same coordinates (see zoom-in).

Figure 12 shows the SO<sub>2</sub> LH retrieved by IASI from 27 until 29th of June. Note that the IASI overpass is around 02:00h UTC and hence about 6 hours after the TROPOMI overpass. Nevertheless the LHs retrieved close to the volcano are in very good agreement with the FP\_ILM results, being also at about 4–6 km. Similarly for the extended plume, for which only a few pixels are shown in Fig. 10, the agreement is very good, with LHs of about 13 km.

### 4.3 Raikoke

On 22 June 2019 at 04:00h local time, the Raikoke stratovolcano located on the Kuril Islands (Russia, 551 m summit elevation) erupted explosively (VEI ≥ 4) after being dormant since 1924. There were several strong distinct explosions, producing a dense ash and SO<sub>2</sub> plume that rose until 13 km altitude the first days and was entrained into the stratosphere (see Sennert 2019). This was the strongest volcanic eruption since the Merapi eruption in 2011, producing a colossal SO<sub>2</sub> plume with an SO<sub>2</sub> loading of more than 900 DU on 22 June 2019, that was dispersed by strong winds over Russia and North America, and was detectable even two weeks after the volcanic eruption. The Raikoke eruption was still ongoing during the writing of this paper. Further results for this eruption are expected to be subject of a future publication.

Figure 13 shows SO<sub>2</sub> VCD measured by TROPOMI for the first three days after the eruption. Note that the SO<sub>2</sub> plume is close to the date line with several overlapping S5P orbits at different overpass times. Therefore we have chosen to plot only one single orbit per day in the images. The figure shows the plume at overpass times of around 00:00h and 02:00h UTC.

Figure 14 shows the corresponding SO<sub>2</sub> LH retrieved by the FP\_ILM for SO<sub>2</sub> VCD ≥ 20 DU. Clearly, the plume shows several layers, with SO<sub>2</sub> LH ranging from 6–8 km up to 19 km on 23 June and from 11 km up to 20 km on 24 June 2019. This is in very good agreement with MLS data (colored circles in the figure) with overpass times around 02:20h UTC on 23 June (left plot), around 01:30h UTC on 24 June (center plot) and around 00:30 UTC on 25 June (left plot): The MLS profiles in Fig. 15 show an SO<sub>2</sub> layer at 17 km on 23 June, two distinct layers at 12 and 17 km at different positions of the plume on 24 June and a layer in the range from 12–14 km on 25 June.

In addition, CALIOP/CALIPSO was able to detect the ash plume during the first days after the eruption, with plume height values in very good agreement with the FP\_ILM SO<sub>2</sub> LH results. On 22 June an ash plume at 17 km was detected (not shown here), The CALIPSO data from 23 June (overpass around 01:30h UTC) shown in Fig. 16 (left) show an ash layer around 5–8 km in the northern part of the plume (i.e. around 46°N), and a second ash layer around 13 km in the southern part, in very

good agreement with the FP\_ILM SO<sub>2</sub> LH results. On 24 June an ash layer at 13 km is visible which is in agreement with the lowest SO<sub>2</sub> LHs retrieved. Note that the CALIPSO ground tracks are shown in Fig. 14 (blue line).

The latest IASI retrievals indicate that the bulk of the SO<sub>2</sub> mass is contained within 9–15 km (not shown here). We note that at the time of writing this paper, the IASI LH retrieval algorithm is undergoing improvements to handle cases with very large SO<sub>2</sub> loadings in the center of the plume. Since the Raikoke eruption was still ongoing during the writing of this paper, we refer here to an upcoming publication of the IASI results for the Raikoke eruption.

## 5 Implementation in an operational environment

In this section we describe the manner in which the algorithm could be implemented in the operational S5P/TROPOMI ground segment.

- 10 Since the same input parameters are used for retrieving operational SO<sub>2</sub> VCD, i.e. O<sub>3</sub> VCD, viewing parameters and surface conditions as well as radiance and irradiance data, the FP\_ILM algorithm can be easily integrated within the operational UPAS processor used for generating the SO<sub>2</sub> products. In the operational TROPOMI environment, cloud properties and the O<sub>3</sub> VCD are retrieved before the SO<sub>2</sub> algorithm is started. Thus all required input parameters are already available when the FP\_ILM algorithm is triggered in the case of a volcanic eruption. Only the wavelength calibration and calculation of the reflectance spectrum has to be performed prior to retrieving the SO<sub>2</sub> LH, since this step differs from the operational SO<sub>2</sub> VCD retrieval.

To process a satellite pixel containing volcanic SO<sub>2</sub>, the FP\_ILM algorithm should be triggered by the operational enhanced SO<sub>2</sub> detection flag (see Theys 2018 for a reference) and using a threshold (a priori) SO<sub>2</sub> VCD of at least 10 DU. The resulting SO<sub>2</sub> LH for this pixel should be stored in the final L2 SO<sub>2</sub> product and can be also used to calculate an optimized SO<sub>2</sub> VCD for this LH.

- 20 The SO<sub>2</sub> LH retrieval only takes about 2 ms per TROPOMI spectrum, hence even for an extended volcanic plume, the entire LH retrieval can be performed in a matter of seconds. This is important for operational retrieval environments with strong time constraints. Currently, the entire processing of a single TROPOMI pixel (i.e. cloud parameters, O<sub>3</sub> and SO<sub>2</sub> VCD retrieval) in the operational ground segment takes about 90 ms, which translates to about 24 min for an entire S5P orbit. Thus the additional time required for the retrieval of the SO<sub>2</sub> LH for selected pixels is not significant.

## 25 6 Conclusions

- We have developed a new algorithm for the fast and accurate retrieval of SO<sub>2</sub> layer heights from UV earthshine observations of volcanic SO<sub>2</sub> eruptions by the TROPOMI sensor on board the Sentinel-5 Precursor platform. The SO<sub>2</sub> LH retrieval has two phases - (1) a computationally expensive off-line training phase in which the retrieval inverse operator is obtained using the FP\_ILM (Full Physics Learning Machine) algorithm; and (2) a fast operational phase, in which the FP\_ILM inverse operator is applied to measured UV reflectance spectra. The FP\_ILM combines a Principal Component Analysis with a Neural Network regression using the UV reflectance, O<sub>3</sub> total column, viewing geometries and surface properties as input. Based on an

optimized training dataset created with smart sampling techniques, the principal component scores calculated for reflectance spectra in the wavelength range 310-335 nm are used along with the other parameters to train a feed-forward artificial neural network. For S5P/TROPOMI measurement data, an initial dimensionality reduction of the reflectance spectra is performed by applying the PCA-derived Principal Component scores before retrieving SO<sub>2</sub> LH with the trained Neural Network.

- 5 The FP\_ILM can be used for NRT applications with strict time constraints. S5P/TROPOMI, with its high spatial and spectral resolution, provides a huge amount of data and the computationally intensive direct fitting approaches to SO<sub>2</sub> LH retrieval developed so far are not applicable. In contrast, the FP\_ILM operator performs SO<sub>2</sub> LH retrieval within about 2 ms per TROPOMI spectrum. Hence even the retrieval on extended volcanic plumes can be performed in a matter of few seconds, allowing for the determination of the SO<sub>2</sub> layer height with an accuracy better than 2 km for SO<sub>2</sub> total column densities larger than 20 DU.
- 10 In this paper, we deployed an independent simulated reflectance spectra dataset to investigate the accuracy of SO<sub>2</sub> LH retrievals and their dependencies on a number of different factors and parameters both direct and indirect. In particular, it was found that retrieved SO<sub>2</sub> LH is strongly dependent on the SO<sub>2</sub> VCD for low VCDs  $\leq 20$  DU as well as for high SZAs  $\geq 75^\circ$ . For high VCDs and low SZA, **according to closed-loop tests**, the SO<sub>2</sub> LH can be retrieved with an accuracy of better than 2 km. We also investigated the dependencies on ISRF and SNR, both of which turned out to be relatively slight effects.
- 15 Broad-band spectral scattering and absorption due to sulfate aerosols or volcanic ash plumes will certainly affect SO<sub>2</sub> LH retrievals. Although not considered in the present work, we will address this important issue in a forthcoming paper in this series; aerosols will be accounted for explicitly in the training of the FP\_ILM. Nevertheless, we should note that SO<sub>2</sub> and ash are likely to be collocated only for fresh volcanic plumes. For mature plumes, mass differences will ensure that ash and SO<sub>2</sub> plumes are not located at similar altitudes, and the corresponding plumes are thus subject to different wind-direction dispersal.
- 20 We have applied the FP\_ILM to a number of volcanic eruptions (**VEI $\geq 3$** ) observed recently by S5P/TROPOMI. Our SO<sub>2</sub> LH results have been compared to SO<sub>2</sub> LHs retrieved from IR measurements from two MetOp/IASI satellites, **as well as NASA AURA/MLS SO<sub>2</sub> profile and CALIOP/CALIPSO ash and aerosol measurements**. Unfortunately the orbits of the MetOp satellites and S5P have widely different overpass times, allowing only for qualitative comparisons. Despite this, there is in general a very good agreement between the IASI and TROPOMI results. **Very good agreement with MLS data was found,**
- 25 **which provides SO<sub>2</sub> profiles with limited spatial coverage and vertical resolution, however with overpass times close to the S5P overpass. Nevertheless,** further verification work is certainly needed. For the Sinabung eruption in February 2018 **and the Raikoke eruption in 2019, our results also showed excellent agreement with CALIOP/CALIPSO LIDAR data measuring the ash plume.**

- Acknowledgements.* We are grateful to Nicolas Theys for helpful discussions of the SO<sub>2</sub> layer height retrieval from UV sensors. We hereby
- 30 acknowledge financial support from DLR programmatic (S5P KTR 2472046) for the development of TROPOMI retrieval algorithms. **We acknowledge ESA and KNMI for the Sentinel-5 precursor L1 data and NASA for the CALIPSO and MLS data used in the paper. We would like to thank two anonymous referees for their comments that helped us improve the manuscript.** L.C. is a research associate supported by the Belgian F.R.S.-FNRS.

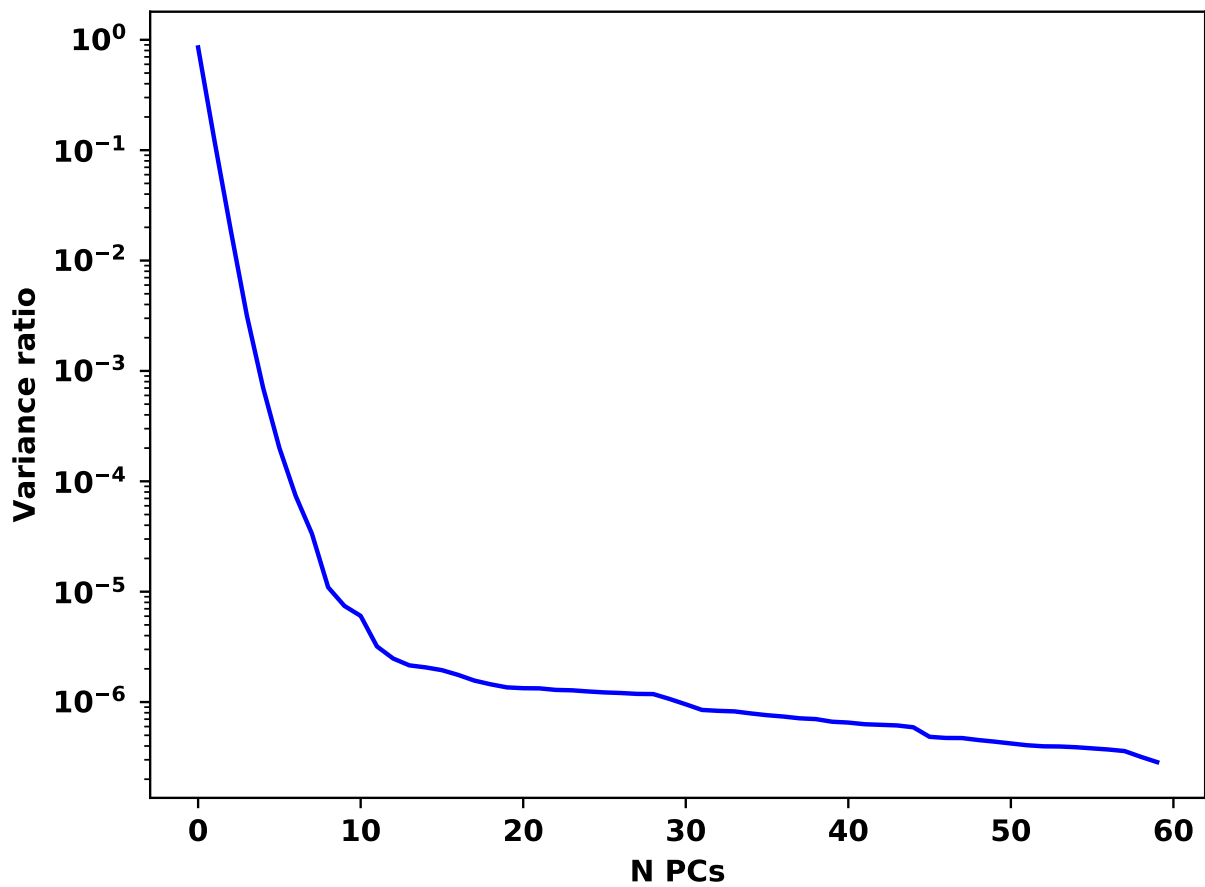


## References

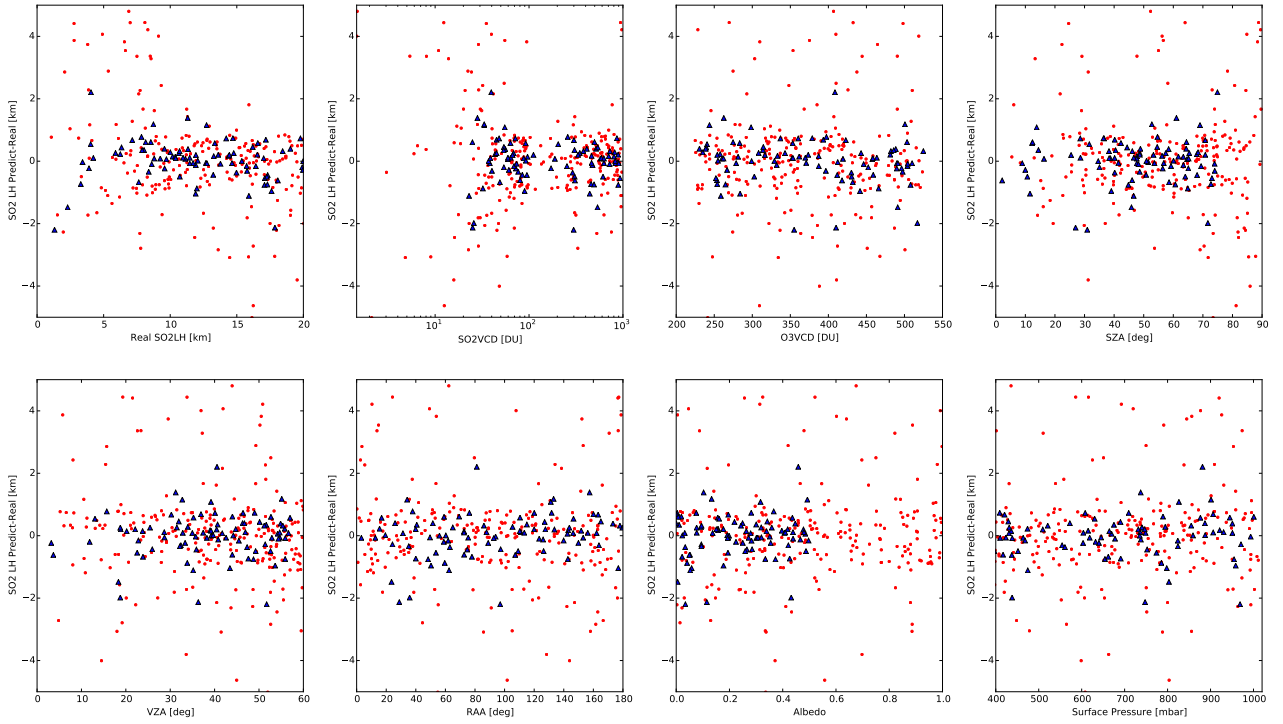
- Bhartia, P.: Algorithm Theoretical Baseline Document, TOMS v8 Total ozone algorithm, Tech. rep., NASA, Greenbelt, Md., [http://toms.gsfc.nasa.gov/version8/version8\\_update.html](http://toms.gsfc.nasa.gov/version8/version8_update.html), 2003.
- Carboni, E., Grainger, R., Walker, J., Dudhia, A., and Siddans, R.: A new scheme for sulphur dioxide retrieval from IASI measurements: application to the Eyjafjallajökull eruption of April and May 2010, *Atmospheric Chemistry and Physics*, 12, 11 417–11 434, <https://doi.org/10.5194/acp-12-11417-2012>, <https://www.atmos-chem-phys.net/12/11417/2012/>, 2012.
- Carboni, E., Grainger, R. G., Mather, T. A., Pyle, D. M., Thomas, G. E., Siddans, R., Smith, A. J. A., Dudhia, A., Koukouli, M. E., and Balis, D.: The vertical distribution of volcanic SO<sub>2</sub> plumes measured by IASI, *Atmospheric Chemistry and Physics*, 16, 4343–4367, <https://doi.org/10.5194/acp-16-4343-2016>, <https://www.atmos-chem-phys.net/16/4343/2016/>, 2016.
- 10 Clarisse, L., Coheur, P. F., Prata, A. J., Hurtmans, D., Razavi, A., Phulpin, T., Hadji-Lazaro, J., and Clerbaux, C.: Tracking and quantifying volcanic SO<sub>2</sub> with IASI, the September 2007 eruption at Jebel at Tair, *Atmospheric Chemistry & Physics*, 8, 7723–7734, 2008.
- Clarisse, L., Coheur, P.-F., Theys, N., Hurtmans, D., and Clerbaux, C.: The 2011 Nabro eruption, a SO<sub>2</sub> plume height analysis using IASI measurements, *Atmospheric Chemistry & Physics*, 14, 3095–3111, <https://doi.org/10.5194/acp-14-3095-2014>, 2014.
- Efremenko, D. S., Loyola R., D. G., Hedelt, P., and Spurr, R. J. D.: Volcanic SO<sub>2</sub> plume height retrieval from UV sensors using a full-physics inverse learning machine algorithm, *International Journal of Remote Sensing*, 38, 1–27, <https://doi.org/10.1080/01431161.2017.1348644>, 2017.
- 20 Fioletov, V. E., McLinden, C. A., Krotkov, N., Yang, K., Loyola, D. G., Valks, P., Theys, N., Van Roozendaal, M., Nowlan, C. R., Chance, K., Liu, X., Lee, C., and Martin, R. V.: Application of OMI, SCIAMACHY, and GOME-2 satellite SO<sub>2</sub> retrievals for detection of large emission sources, *Journal of Geophysical Research: Atmospheres*, 118, 11,399–11,418, <https://doi.org/10.1002/jgrd.50826>, <http://dx.doi.org/10.1002/jgrd.50826>, 2013JD019967, 2013.
- Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, J., Betts, R., Fahey, D. W., Haywood, J., Lean, J., Lowe, D. C., Myhre, G., Nganga, J., Prinn, R., Raga, G., Schulz, M., van Dorland, R., Bodeker, G., Boucher, O., Collins, W. D., Conway, T. J., Dlugokencky, E., Elkins, J. W., Etheridge, D., Foukal, P., Fraser, P., Geller, M., Joos, F., Keeling, C. D., Keeling, R., Kinne, S., Lassey, K., Lohmann, U., Manning, A. C., Montzka, S. A., Oram, D., O’Shaughnessy, K., Piper, S. C., Plattner, G. K., Ponater, M., Ramankutty, N., Reid, G. C., Rind, D., Rosenlof, K. H., Sausen, R., Schwarzkopf, D., Solanki, S. K., Stenchikov, G., Stuber, N., Takemura, T., Textor, C., Wang, R., Weiss, R., and Whorf, T.: Changes in atmospheric constituents and in radiative forcing, chap. 2, pp. 129–234, Cambridge University Press, Cambridge; New York, n/a, 2007.
- 25 Kleipool, Q., Ludewig, A., Babić, L., Bartstra, R., Braak, R., Dierssen, W., Dewitte, P.-J., Kenter, P., Landzaat, R., Leloux, J., Loots, E., Meijering, P., van der Plas, E., Rozemeijer, N., Schepers, D., Schiavini, D., Smeets, J., Vacanti, G., Vonk, F., and Veeffkind, P.: Pre-launch calibration results of the TROPOMI payload on-board the Sentinel-5 Precursor satellite, *Atmospheric Measurement Techniques*, 11, 6439–6479, <https://doi.org/10.5194/amt-11-6439-2018>, <https://www.atmos-meas-tech.net/11/6439/2018/>, 2018.
- 30 Krueger, A. J., Walter, L. S., Bhartia, P. K., Schnetzler, C. C., Krotkov, N. A., Sprod, I., and Bluth, G. J. S.: Volcanic sulfur dioxide measurements from the total ozone mapping spectrometer instruments, *Journal of Geophysical Research*, 100, 14,057–14,076, <https://doi.org/10.1029/95JD01222>, 1995.
- 35 Lee, C., Martin, R. V., van Donkelaar, A., Lee, H., Dickerson, R. R., Hains, J. C., Krotkov, N., Richter, A., Vinnikov, K., and Schwab, J. J.: SO<sub>2</sub> emissions and lifetimes: Estimates from inverse modeling using in situ and global, space-based (SCIAMACHY and OMI)

- observations, *Journal of Geophysical Research: Atmospheres*, 116, n/a–n/a, <https://doi.org/10.1029/2010JD014758>, <http://dx.doi.org/10.1029/2010JD014758>, d06304, 2011.
- Lerot, C., Van Roozendaal, M., Spurr, R., Loyola, D., Coldewey-Egbers, M., Kochenova, S., van Gent, J., Koukouli, M., Balis, D., Lambert, J.-C., Granville, J., and Zehner, C.: Homogenized total ozone data records from the European sensors GOME/ERS-2, SCIAMACHY/Envisat, and GOME-2/MetOp-A, *Journal of Geophysical Research: Atmospheres*, 119, 1639–1662, <https://doi.org/10.1002/2013JD020831>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD020831>, 2014.
- Li, C., Krotkov, N. A., Carn, S., Zhang, Y., Spurr, R. J. D., and Joiner, J.: New-generation NASA Aura Ozone Monitoring Instrument (OMI) volcanic SO<sub>2</sub> dataset: algorithm description, initial results, and continuation with the Suomi-NPP Ozone Mapping and Profiler Suite (OMPS), *Atmospheric Measurement Techniques*, 10, 445–458, <https://doi.org/10.5194/amt-10-445-2017>, <https://www.atmos-meas-tech.net/10/445/2017/>, 2017.
- Loyola, D., Xu, J. Heue, K. P., and Zimmer, W.: Applying FP\_ILM to the retrieval of geometry-dependent effective Lambertian equivalent reflectivity (GE\_LER) to account for BRDF effects on UVN satellite measurements of trace gases, clouds and aerosols, *Atmos. Meas. Tech. Discuss.*, in review, <https://doi.org/10.5194/amt-2019-37>, 2019.
- Loyola, D. G.: Applications of neural network methods to the processing of earth observation satellite data, *Neural Networks*, 19, 168–177, <https://doi.org/10.1016/j.neunet.2006.01.010>, <http://dx.doi.org/10.1016/j.neunet.2006.01.010>, 2006.
- Loyola, D. G., Pedernana, M., and Gimeno Garcia, S.: Smart sampling and incremental function learning for very large high dimensional data, *Neural Networks*, 78, 75–87, <https://doi.org/10.1016/j.neunet.2015.09.001>, <http://dx.doi.org/10.1016/j.neunet.2015.09.001>, 2016.
- Malavelle, F. F., Haywood, J. M., Jones, A., Gettelman, A., Clarisse, L., Bauduin, S., Allan, R. P., Karset, I. H. H., Kristjánsson, J. E., Oreopoulos, L., Cho, N., Lee, D., Bellouin, N., Boucher, O., Grosvenor, D. P., Carslaw, K. S., Dhomse, S., Mann, G. W., Schmidt, A., Coe, H., Hartley, M. E., Dalvi, M., Hill, A. A., Johnson, B. T., Johnson, C. E., Knight, J. R., O'Connor, F. M., Partridge, D. G., Stier, P., Myhre, G., Platnick, S., Stephens, G. L., Takahashi, H., and Thordarson, T.: Strong constraints on aerosol-cloud interactions from volcanic eruptions, *Nature*, 546, 485–491, <https://doi.org/10.1038/nature22974>, 2017.
- McCormick, M. P., Thomason, L. W., and Trepte, C. R.: Atmospheric effects of the Mt Pinatubo eruption, *Nature*, 373, 399–404, <https://doi.org/10.1038/373399a0>, 1995.
- Myles, L., Meyers, T. P., and Robinson, L.: Relaxed eddy accumulation measurements of ammonia, nitric acid, sulfur dioxide and particulate sulfate dry deposition near Tampa, FL, USA, *Environmental Research Letters*, 2, <https://doi.org/10.1088/1748-9326/2/3/034004>, 2011.
- Nowlan, C. R., Liu, X., Chance, K., Cai, Z., Kurosu, T. P., Lee, C., and Martin, R. V.: Retrievals of sulfur dioxide from the Global Ozone Monitoring Experiment 2 (GOME-2) using an optimal estimation approach: Algorithm and initial validation, *Journal of Geophysical Research: Atmospheres*, 116, n/a–n/a, <https://doi.org/10.1029/2011JD015808>, <http://dx.doi.org/10.1029/2011JD015808>, d18301, 2011.
- Pedernana, M., Loyola, D., Apituley, A., Sneep, M., and Veefkind, P.: Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Sulphur Dioxide SO<sub>2</sub>, Tech. Rep. 1.1.2, BIRA, <https://sentinel.esa.int/documents/247904/2474726/Sentinel-5P-Level-2-Product-User-Manual-Sulphur-Dioxide>, 2018.
- Pumphrey, H. C., Read, W. G., Livesey, N. J., and Yang, K.: Observations of volcanic SO<sub>2</sub> from MLS on Aura, *Atmospheric Measurement Techniques*, 8, 195–209, <https://doi.org/10.5194/amt-8-195-2015>, <https://www.atmos-meas-tech.net/8/195/2015/>, 2015.
- Rix, M., Valks, P., Hao, N., Loyola, D., Schlager, H., Huntrieser, H., and Flemming, J. and Koehler, U. and Schumann, U. and Inness, A.: Volcanic SO<sub>2</sub>, BrO and plume height estimations using GOME-2 satellite measurements during the eruption of Eyjafjallajökull in May 2010, *Journal of Geophysical Research (Atmospheres)*, 117, D00U19, <https://doi.org/10.1029/2011JD016718>, 2012.
- Robock, A.: Volcanic eruptions and climate, *Reviews of Geophysics*, 38, 191–219, <https://doi.org/10.1029/1998RG000054>, 2000.

- Sennert, S. K., ed.: Report on Raikoke (Russia), Weekly Volcanic Activity Report, 19 June-25 June 2019, Global Volcanism Program, Smithsonian Institution and US Geological Survey, 2019.
- Spurr, R., de Haan, J., van Oss, R., and Vasilkov, A.: Discrete-ordinate radiative transfer in a stratified medium with first-order rotational Raman scattering, *Journal of Quantitative Spectroscopy & Radiative Transfer*, 109, 404–425, <https://doi.org/10.1016/j.jqsrt.2007.08.011>, 2008.
- Theys, N., De Smedt, I., Yu, H., Danckaert, T., van Gent, J., Hörmann, C., Wagner, T., Hedelt, P., Bauer, H., Romahn, F., Pedernana, M., Loyola, D., and Van Roozendaal, M.: Sulfur dioxide retrievals from TROPOMI onboard Sentinel-5 Precursor: algorithm theoretical basis, *Atmospheric Measurement Techniques*, 10, 119–153, <https://doi.org/10.5194/amt-10-119-2017>, <http://www.atmos-meas-tech.net/10/119/2017/>, 2017.
- 10 Theys, N. e. a.: S5P/TROPOMI SO<sub>2</sub> ATBD, Tech. Rep. 1.1.0, BIRA, <https://sentinels.copernicus.eu/documents/247904/2476257/Sentinel-5P-ATBD-SO2-TROPOMI>, 2018.
- Venzke, E., ed.: Report on Sierra Negra (Ecuador), Bulletin of the Global Volcanism Network, vol. 43:9 of *Global Volcanism Program*, Smithsonian Institution, 2018a.
- Venzke, E., ed.: Report on Sinabung (Indonesia), Bulletin of the Global Volcanism Network, vol. 43:4 of *Global Volcanism Program*, 15 Smithsonian Institution, 2018b.
- von Glasow, R., Bobrowski, N., and Kern, C.: The effects of volcanic eruptions on atmospheric chemistry, *Chemical Geology*, 263, 131 – 142, <https://doi.org/10.1016/j.chemgeo.2008.08.020>, halogens in Volcanic Systems and Their Environmental Impacts, 2009.
- Winker, D. M., Pelon, J., Coakley, J. A., Ackerman, S. A., Charlson, R. J., Colarco, P. R., Flamant, P., Fu, Q., Hoff, R. M., Kittaka, C., Kubar, T. L., Le Treut, H., McCormick, M. P., MÃ©gie, G., Poole, L., Powell, K., Trepte, C., Vaughan, M. A., and Wielicki, B. A.: The CALIPSO Mission, *Bulletin of the American Meteorological Society*, 91, 1211–1230, <https://doi.org/10.1175/2010BAMS3009.1>, 2010.
- 20 Xu, J., Schüssler, O., Loyola R., D., Romahn, F., and Doicu, A.: A novel ozone profile shape retrieval using Full-Physics Inverse Learning Machine (FP-ILM)., *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 10, 5442–5457, <https://doi.org/10.1109/JSTARS.2017.2740168>, 2017.
- Yang, K., Liu, X., Krotkov, N. A., Krueger, A. J., and Carn, S. A.: Estimating the altitude of volcanic sulfur dioxide plumes from space borne hyper-spectral UV measurements, *Geophysical Research Letters*, 36, <https://doi.org/10.1029/2009GL038025>, <http://dx.doi.org/10.1029/2009GL038025>, 2009.
- 25 Yang, K., Liu, X., Bhartia, P. K., Krotkov, N. A., Carn, S. A., Hughes, E. J., Krueger, A. J., Spurr, R. J. D., and Trahan, S. G.: Direct retrieval of sulfur dioxide amount and altitude from spaceborne hyperspectral UV measurements: Theory and application, *Journal of Geophysical Research: Atmospheres*, 115, <https://doi.org/10.1029/2010JD013982>, <http://dx.doi.org/10.1029/2010JD013982>, 2010.



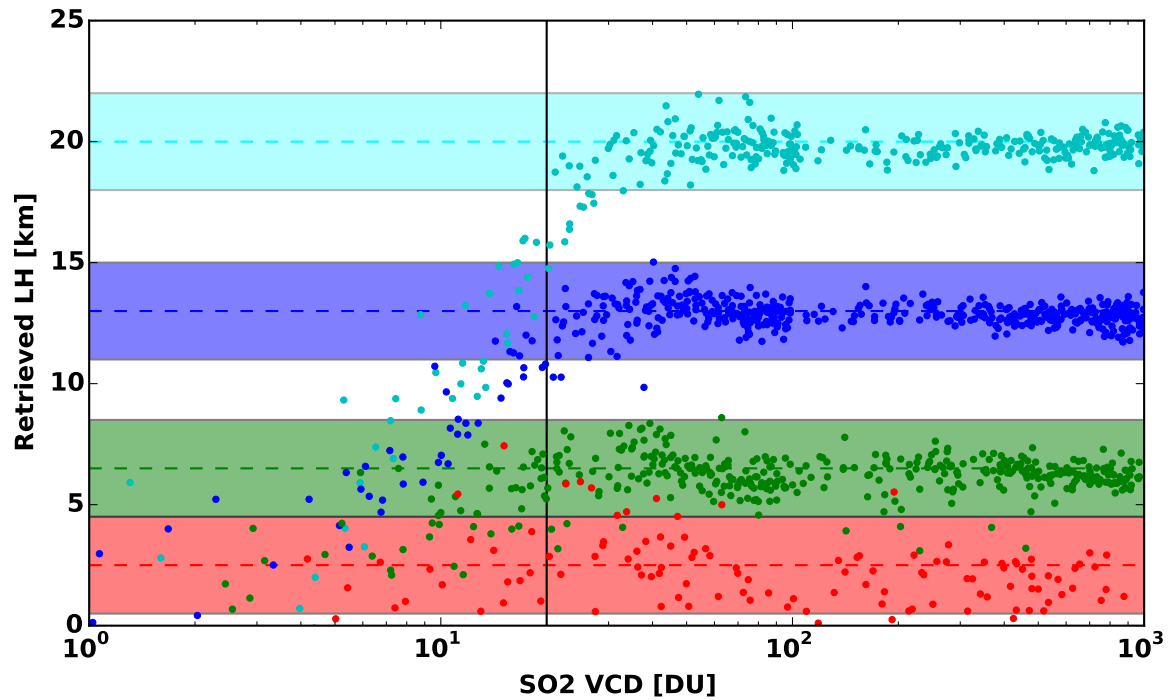
**Figure 1.** [Moved from Figure 6 to Figure 1] "Explained" variance ratio as a function of the number of principal components. Already the inclusion of 10 PCs account for 99.994% of the spectral variance.



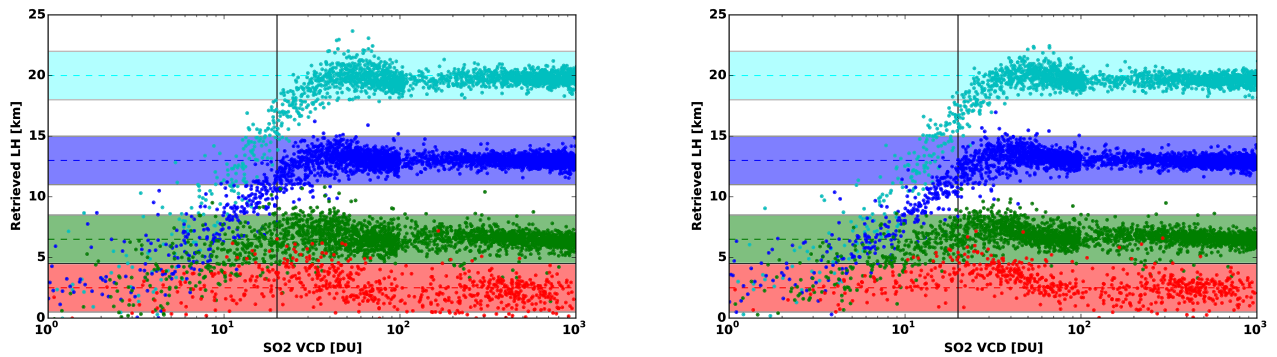
**Figure 2.** [Updated Figure] Dependency of the retrieved SO<sub>2</sub> layer height as a function of 8 parameters as indicated. Plotted are the layer height differences (in [km]) between retrieved layer heights and those simulated using the independent verification dataset (i.e. a 10% subdivision of the entire training dataset). Red dots show the dependencies using the entire training data set, whereas blue triangles show the dependencies with an optimized training dataset (i.e. restricted to SO<sub>2</sub> VCD values  $\geq 20$  DU, SZAs  $\leq 75^\circ$  and albedos  $\leq 0.5$ .)

**Table 1.** Physical parameters varied for the generation of reflectance spectra. The optimized parameter range for training the final FP\_ILM retrieval operator is shown.

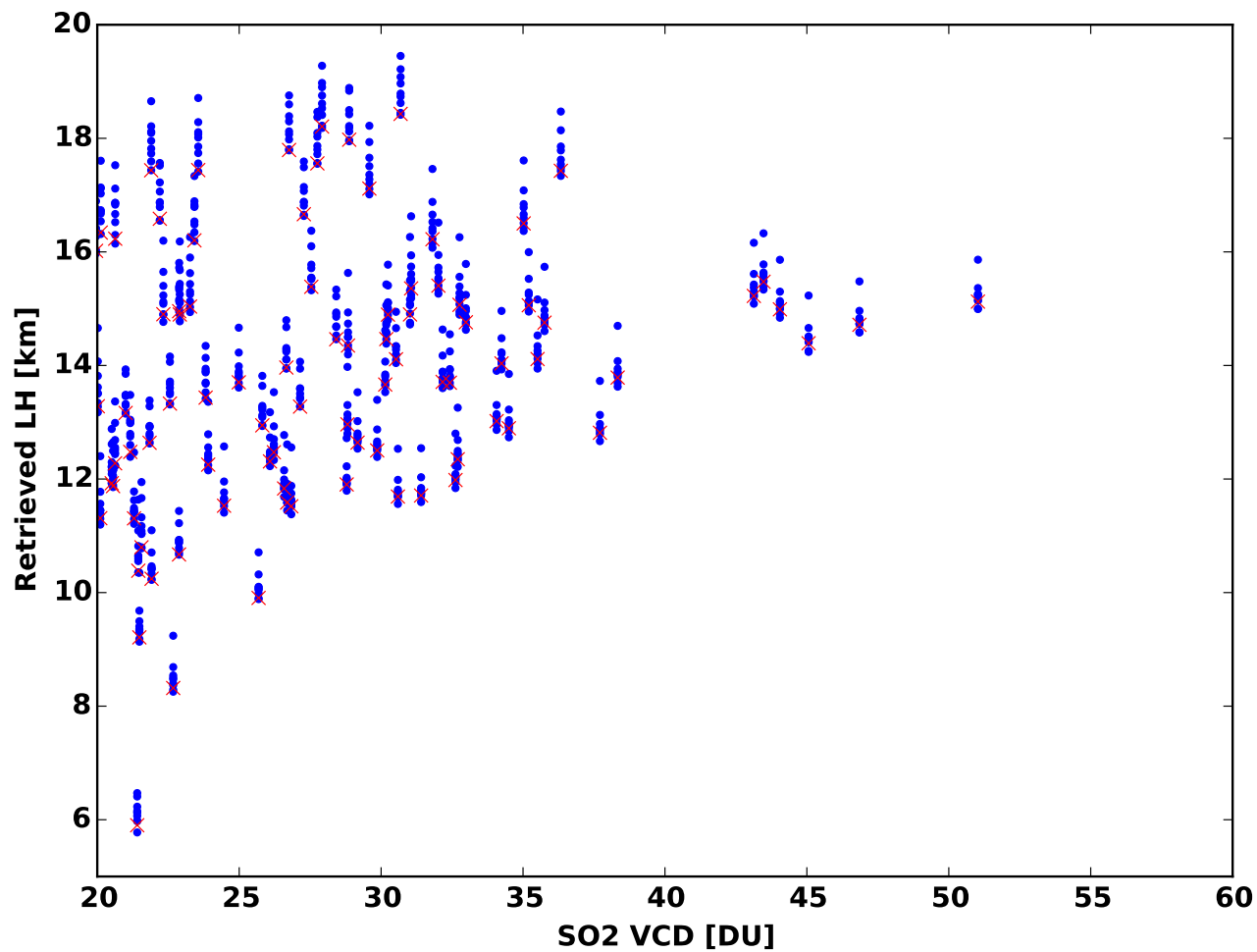
Parameter	Range
SZA	0–75°
VZA	0–75°
RAA	0–180°
Surface albedo	0–0.5
Surface height	0–8 km
O <sub>3</sub> VCD	225–525 DU
SO <sub>2</sub> VCD	0–1000 DU
SO <sub>2</sub> LH	2.5–25 km



**Figure 3.** [Updated Figure] Dependency of the retrieved SO<sub>2</sub> LH as a function of SO<sub>2</sub> total vertical column, using the independent test dataset. Color coded are the different LHs for which the test spectra have been generated. Horizontal color bars indicate a 2 km uncertainty on the retrieved SO<sub>2</sub> LH.

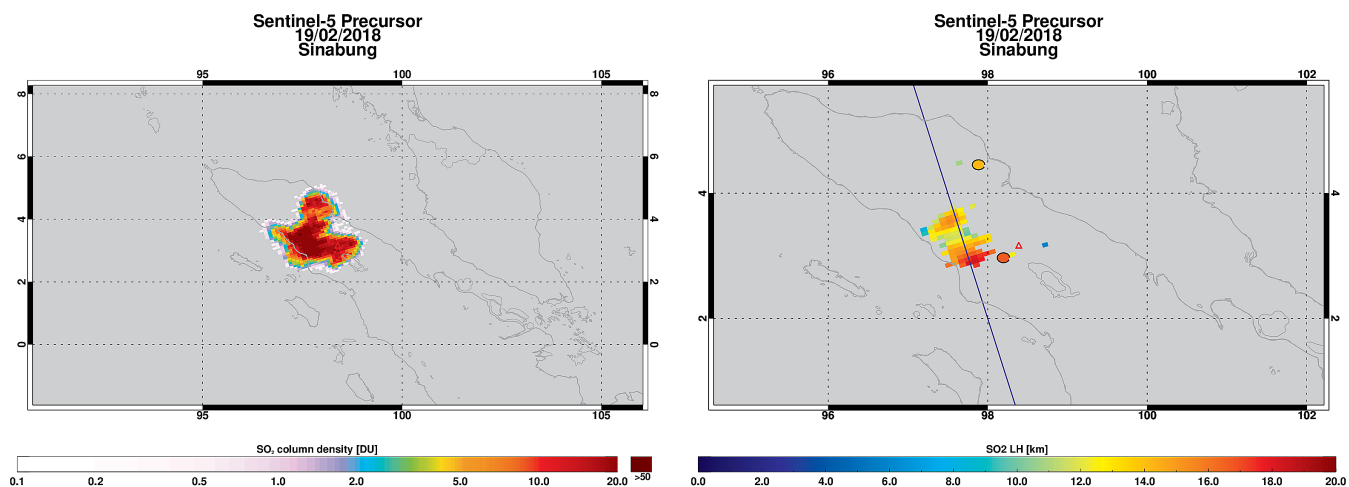


**Figure 4.** [Updated Figure] Dependency of the retrieved SO<sub>2</sub> LH on the SNR. Same as Fig. 3 with different SNR values. Left: SNR=500, Right: 1500.

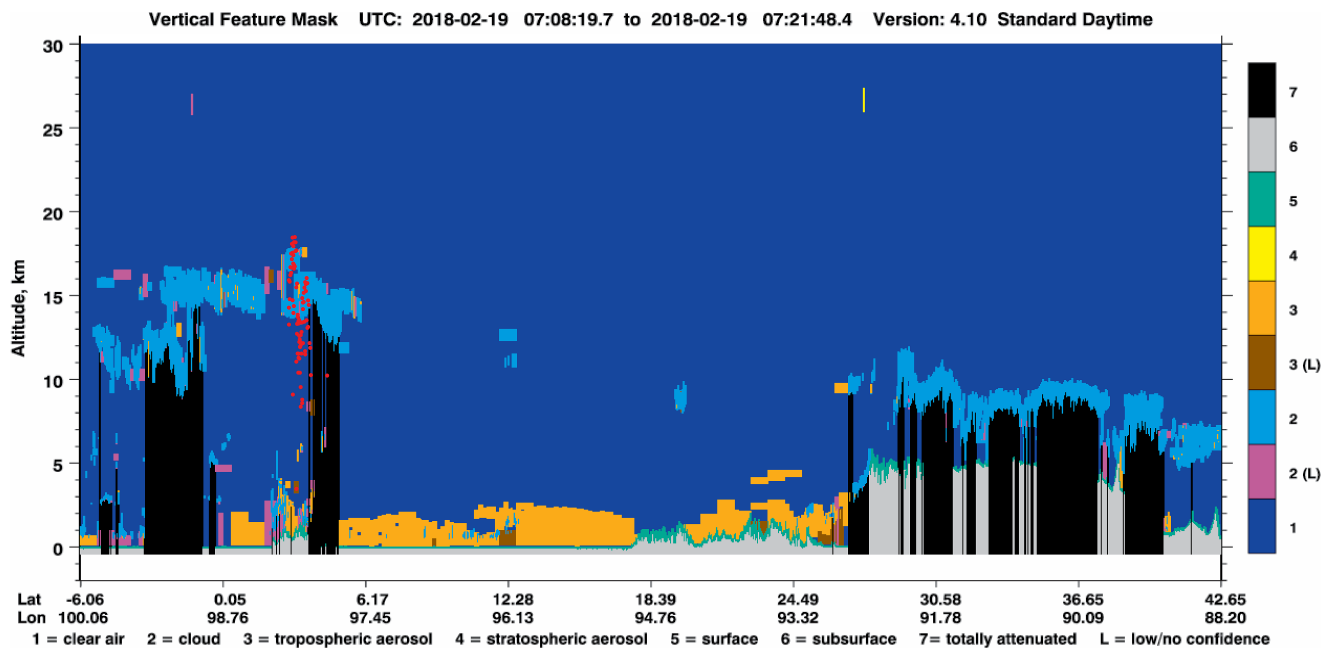


**Figure 5.** Dependency of the retrieved SO<sub>2</sub> LH on the **TROPOMI** instrument row. The SO<sub>2</sub> LH as a function of SO<sub>2</sub> VCD is shown for the Sinabung eruption. Blue dots show the SO<sub>2</sub> LH result for every 50th detector row, whereas red crosses show the LH interpolated to the measurement row.

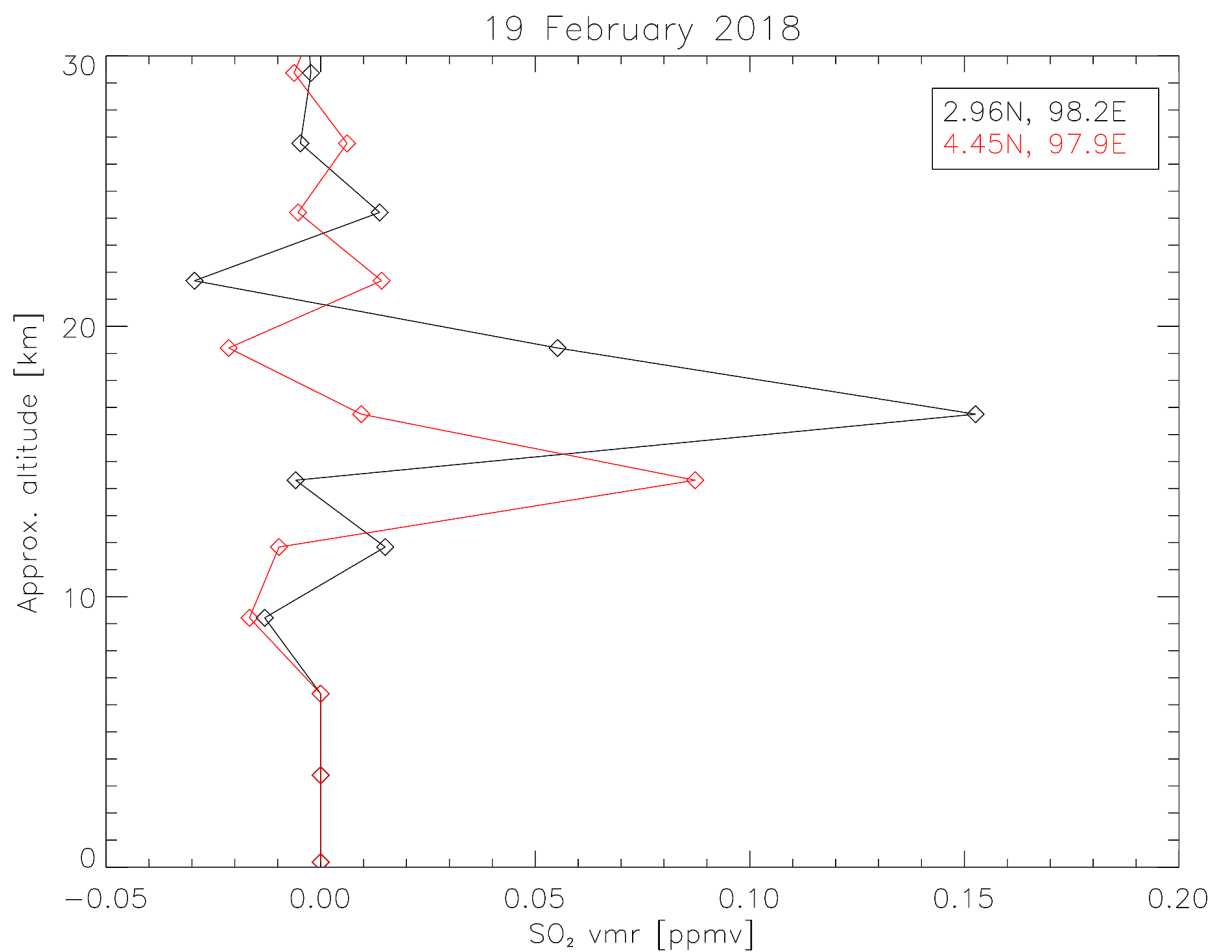




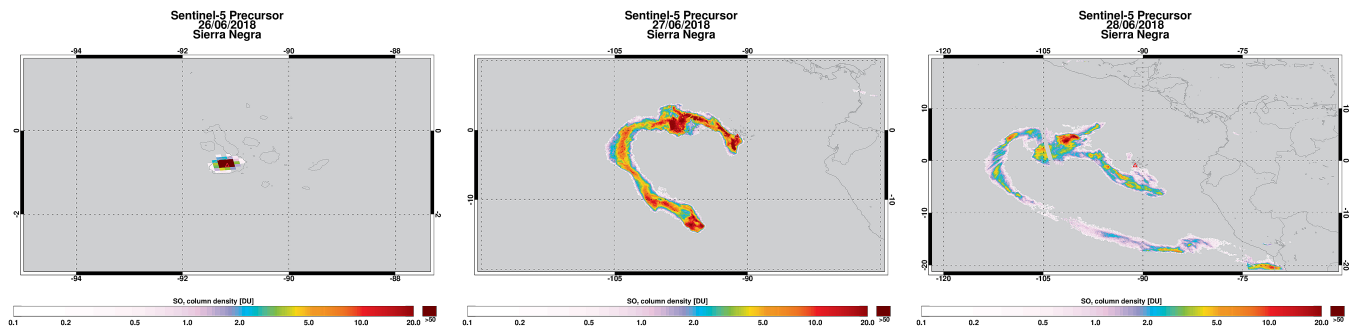
**Figure 6.** [UPDATED FIGURE] SO<sub>2</sub> VCD (left) and SO<sub>2</sub> LH (right) for the TROPOMI measurements of the Sinabung volcano on 19 February 2018, at an overpass time around 06:30h UTC. SO<sub>2</sub> LH results are only shown for VCD  $\geq$  20 DU. In the right plot also MLS/AURA results (the two colored circles) are shown, as well as the CALIPSO groundtrack (cf. Fig. 7)



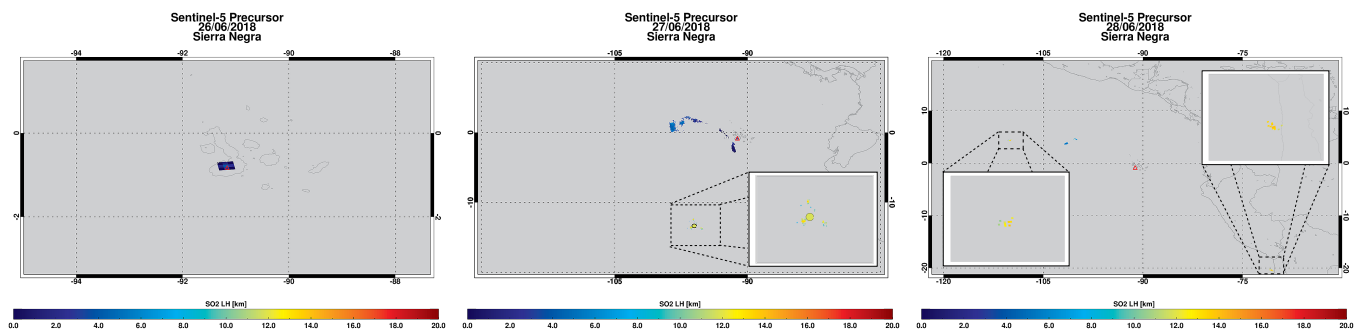
**Figure 7.** [UPDATED FIGURE]CALIOP/CALIPSO vertical feature mask for the measurement of the Sinabung volcanic eruption on 19 February 2018 around 07:15h UTC. Red dots show the retrieved LH for pixels with  $\text{SO}_2 \text{ VCD} \geq 20 \text{ DU}$ . The ground track of the CALIPSO measurements is shown in Fig. 6. The volcano is located at 3.17N, 98.39E. The plume is clearly visible in the left part of the image. Note that the classification scheme sometimes cannot pick up the volcanic ash or sulfate aerosol because of competing clouds. CALIPSO image credit: NASA



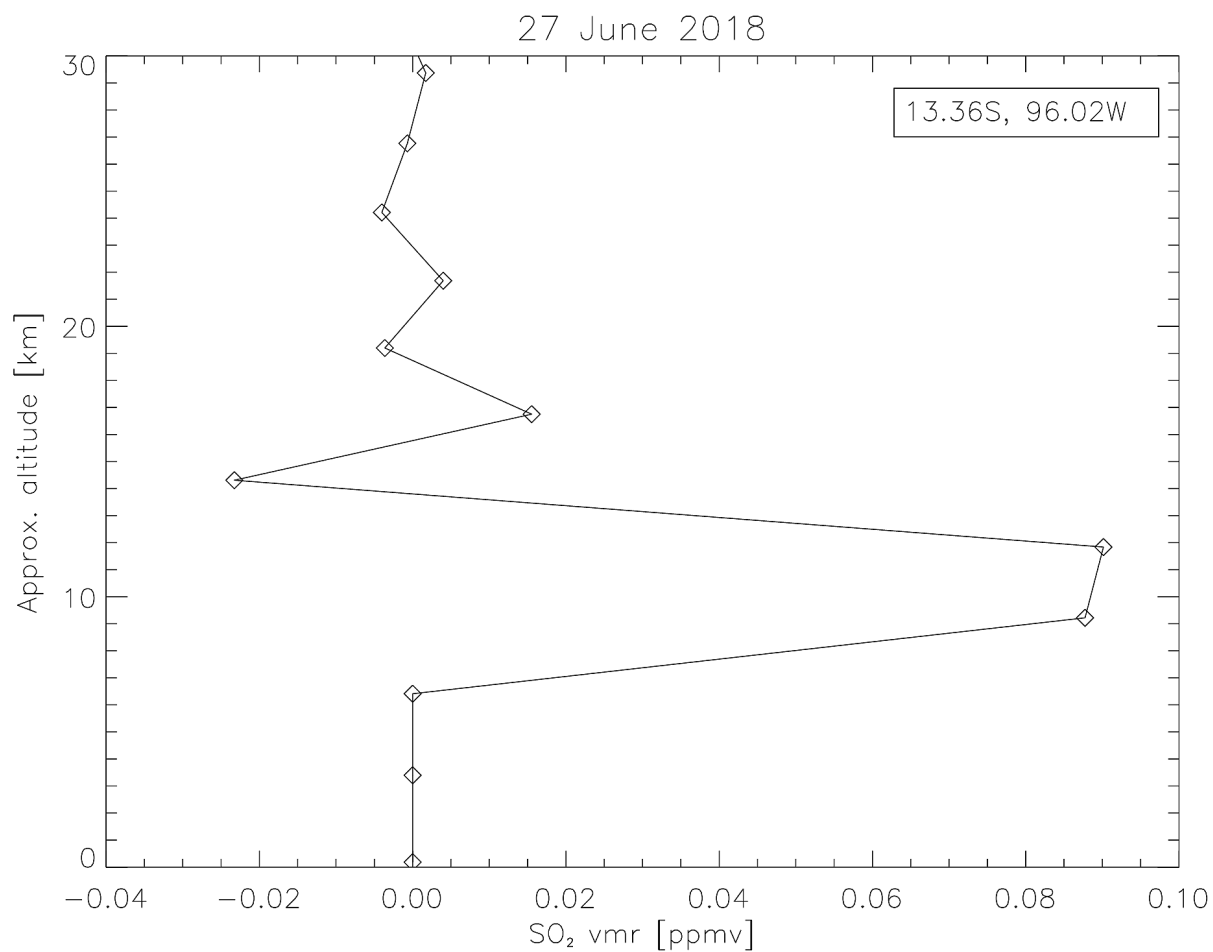
**Figure 8.** [New FIGURE]MLS SO<sub>2</sub> profile for two measurements intersecting the volcanic plume of the Sinabung volcanic eruption on 19 February 2018, clearly showing the presence of an SO<sub>2</sub> layer at an altitude around 16 km close to the volcano (black) and 13 km in the extended plume farther north (red). The position of the measurements is shown in Fig. 6.



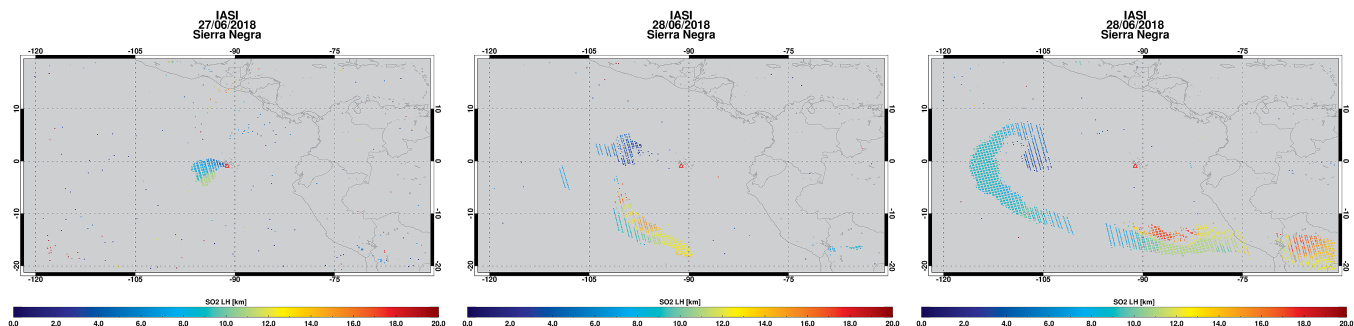
**Figure 9.** [UPDATED FIGURE]SO<sub>2</sub> VCD for the TROPOMI measurements of the Sierra Negra volcano from 26 (left) until 28 June 2018 (right), with overpass times around 19:50h UTC.



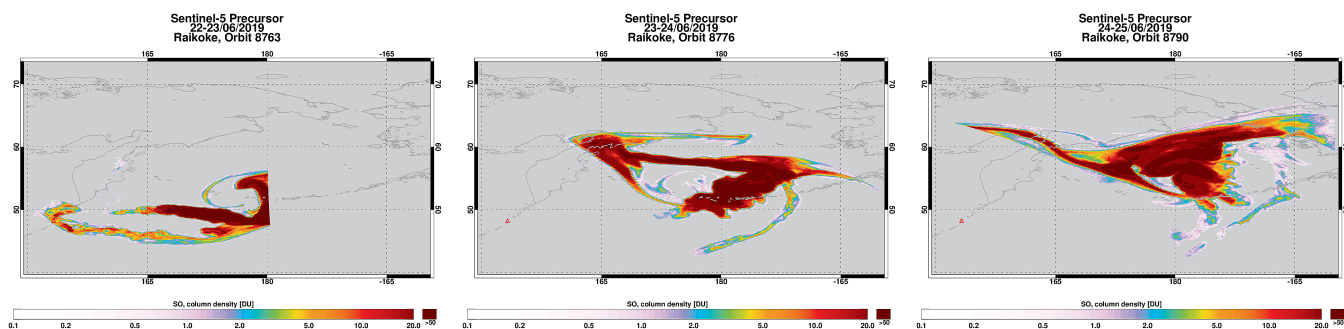
**Figure 10.** [UPDATED FIGURE]FP\_ILM SO<sub>2</sub> LH for the TROPOMI measurements shown in Fig. 9. Note that only pixels with SO<sub>2</sub> VCD $\geq$ 20 DU are shown. The MLS SO<sub>2</sub> measurement of the volcanic plume on 27 June is indicated with a colored circle.



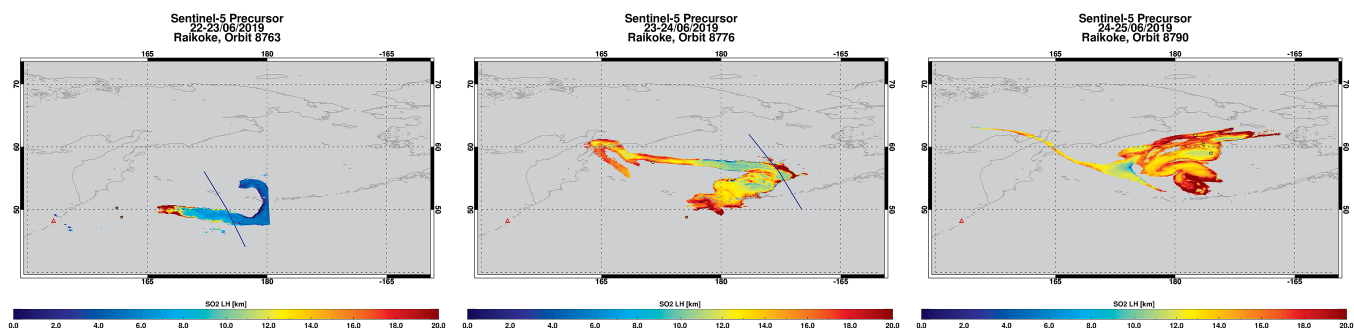
**Figure 11.** [New FIGURE]MLS SO<sub>2</sub> profile measured on 27 February 2018 (at a measurement time of 20:16h UTC) in the extended volcanic plume from the Sierra Negra eruption, clearly showing the presence of an SO<sub>2</sub> layer at an altitude range from 9–11 km. The position of the measurements is shown in Fig. 10.



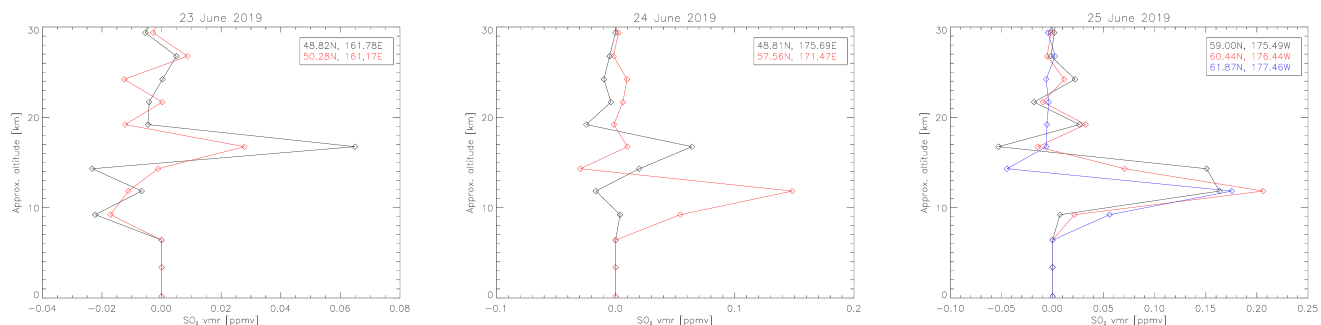
**Figure 12.** [UPDATED FIGURE]MetOp-A and -B IASI SO<sub>2</sub> LH results for the Sierra Negra SO<sub>2</sub> plume from 27 (left) until 29 June 2018 (right). with overpass times around 02:00h, hence about 6 hours after the TROPOMI measurement.



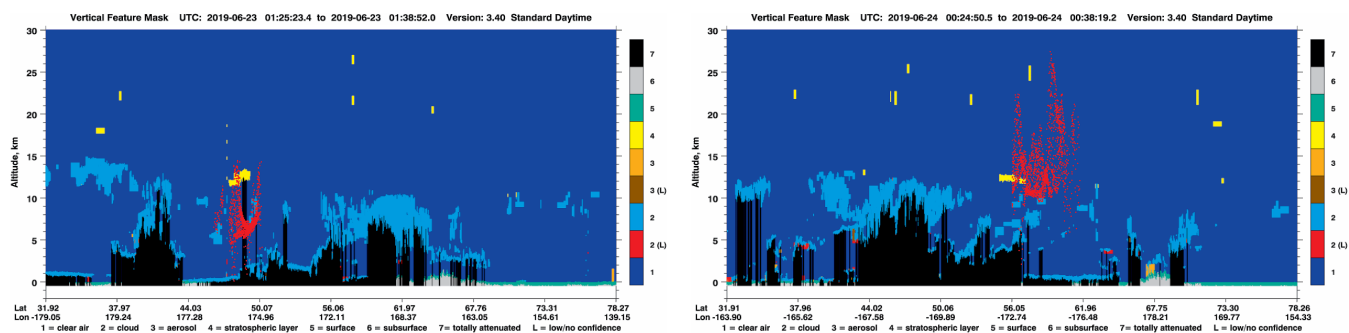
**Figure 13.** [NEW FIGURE]SO<sub>2</sub> VCD for the TROPOMI measurements of the Raikoke volcanic eruption, measured on 22/23 June 2019 (left), 23/24 June 2019 (center) and 24/25 June 2019 (right), with overpass times around 02:00 UTC (orbit 8763, left), and 00:00h UTC (orbits 8776 and 8790).



**Figure 14.** [NEW FIGURE]SO<sub>2</sub> LH retrieved for the TROPOMI orbits shown in Fig. 13. Only pixels with SO<sub>2</sub> VCD ≥ 20 DU are shown. Colored circles indicate the MLS/AURA SO<sub>2</sub> profile peak height for the results shown in Fig. 15. The blue lines indicate the CALIPSO ground track for the results shown in Fig. 16.



**Figure 15.** [NEW FIGURE]MLS SO<sub>2</sub> profiles measured on 23 June 2019 (left, overpass around 02:20h UTC), 24 June 2019 (center, overpass around 01:30h UTC), and 25 June 2019 (right, overpass around 00:30h UTC), in the extended volcanic plume from the Raikoke eruption, clearly showing the presence of an SO<sub>2</sub> layer at an altitude around 17 km on 23 June, two distinct layers at 12 and 17 km at different positions of the plume on 24 June 2019, and around 12–14 km on 25 June 2019. The position of the measurements is shown in Fig. 14.



**Figure 16.** [NEW FIGURE]CALIOP/CALIPSO vertical feature mask for the measurements of the Raikoke plume on 23 June 2019 around 01:30h UTC (left) and on 24 June 2019 around 00:30h UTC (right). Red dots show the retrieved FP\_ILM SO<sub>2</sub> LH for pixels with SO<sub>2</sub> VCD  $\geq$  20 DU. The ground track of the CALIPSO measurements is shown in Fig. 14. Note that the classification scheme sometimes cannot pick up the volcanic ash or sulfate aerosol because of competing clouds. CALIPSO image credit: NASA