

1 Volcanic SO₂ Effective Layer Height Retrieval for OMI Using a 2 Machine Learning Approach

3 Nikita M. Fedkin¹, Can Li², Nickolay A. Krotkov², Pascal Hedelt³, Diego G. Loyola³, Russell R.
4 Dickerson¹, Robert Spurr⁴

5 1: Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA

6 2: NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

7 3: German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), Oberpfaffenhofen, Germany

8 4: RT Solutions Inc., Cambridge, MA, USA

9
10 *Correspondence to:* Nikita M. Fedkin (nfedkin@umd.edu)

11
12 **Abstract.** Information about the height and loading of sulfur dioxide (SO₂) plumes from
13 volcanic eruptions is crucial for aviation safety and for assessing the effect of sulfate aerosols on
14 climate. While SO₂ layer height has been successfully retrieved from backscattered Earthshine
15 ultraviolet (UV) radiances measured by the Ozone Monitoring Instrument (OMI), previously
16 demonstrated techniques are computationally intensive and not suitable for near real-time
17 applications. In this study, we introduce a new OMI algorithm for fast retrievals of effective
18 volcanic SO₂ layer height. We apply the Full Physics Inverse Learning Machine (FP_ILM)
19 algorithm to OMI radiances in the spectral range of 310-330 nm. This approach consists of a
20 training phase that utilizes extensive radiative transfer calculations to generate a large dataset of
21 synthetic radiance spectra for geophysical parameters representing the OMI measurement
22 conditions. The principal components of the spectra from this dataset in addition to a few
23 geophysical parameters are used to train a neural network to solve the inverse problem and
24 predict the SO₂ layer height. This is followed by applying the trained inverse model to real OMI
25 measurements to retrieve the effective SO₂ plume heights. The algorithm has been tested on
26 several major eruptions during the OMI data record. The results for the 2008 Kasatochi, 2014
27 Kelud, 2015 Calbuco, and 2019 Raikoke eruption cases are presented here and compared with
28 volcanic plume heights estimated with other satellite sensors. For the most part, OMI-retrieved
29 effective SO₂ heights agree well with the lidar measurements of aerosol layer height from Cloud-
30 Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) and thermal infrared
31 retrievals of SO₂ heights from the infrared atmospheric sounding interferometer (IASI). The
32 errors in OMI retrieved SO₂ heights are estimated to be 1-1.5 km for plumes with relatively large
33 SO₂ signals (> 40 DU). The algorithm is very fast and retrieves plume height in less than 10 min
34 for an entire OMI orbit.

36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67

1 Introduction

The observation and tracking of emissions from volcanic eruptions are crucial for both air traffic safety and for assessing climate forcing impacts from volcanic sulfate aerosols. In the last 10 years, volcanoes have emitted roughly 20-25 million metric tons of sulfur dioxide (SO₂) per year through passive degassing (Carn et al, 2017). Explosive volcanic eruptions, however, can additionally release large SO₂ amounts high into the atmosphere. SO₂ can be converted to sulfate aerosols within 2-3 days in the troposphere (Lee et al., 2011) and within a few weeks in the lower stratosphere (von Glasow et al., 2009, Krotkov et al., 2010). Sulfate aerosols are known to have a cooling effect on climate, especially if an SO₂ plume is injected into the lower stratosphere and remains there for longer periods of time. This is demonstrated by significant eruptions such as Mt. Pinatubo in 1991 that temporarily reduced global temperatures by up to 0.5°C (McCormick et al, 1995). Aside from releasing SO₂, volcanoes also emit large amounts of ash into the atmosphere which can have adverse impacts on air travel. Ash from volcanic plumes can often interfere with flight paths, greatly reduce visibility near the ground, and cause damage to the aircraft including engine failure (Carn et al., 2009). In addition, SO₂ causes sulfidation in the engines, an effect that can reduce their lifetimes in the long term. From 1953 to 2009, over 120 aviation incidents involving volcanic activity were reported, with roughly 80 of them involving serious damage to the airframe or engine (Guffanti et al., 2010). There is also the possibility of highly concentrated volcanic SO₂ plumes producing acidic aerosols which can cause irritation of the eyes, nose and respiratory airways of occupants inside airplanes (Schmidt et al., 2014). In many cases SO₂ and ash are often collocated, thus making estimates of SO₂ layer height useful for aviation hazard mitigation and volcanic plume forecasting. Lastly, the accurate determination of SO₂ height can ideally aid in producing accurate SO₂ vertical column depth (VCD) estimates given that those retrievals typically use a fixed *a priori* vertical distribution of SO₂ in the absence of additional information on SO₂ height.

With remote sensing, these volcanic plumes can be regularly observed from space. In particular, hyperspectral spectrometers such as the Ozone Monitoring Instrument (OMI), GOME-2, OMPS, TROPOMI and others, have provided frequent and increasingly accurate observations of global SO₂ amounts, through retrieval algorithms from backscattered radiance measurements. The OMI instrument, a Dutch-Finish contribution to the NASA Aura satellite,

68 has been operational since 2004. OMI has 60 cross track positions (rows) and has a $13 \times 24 \text{ km}^2$
69 spatial resolution at the nadir position (Levelt et al., 2006). The instrument uses two UV channels
70 and one visible channel to measure backscattered radiances from the Earth's atmosphere. About
71 half of the OMI rows are affected by the row anomaly which affects the quality of OMI Level 1
72 and Level 2 data. This anomaly affects individual rows and slowly evolves over time. It is
73 thought to occur due to a physical obstruction caused by the loosening of material on the interior
74 of sensor (Torres et al., 2018). In general, SO_2 slant column amounts are retrieved from these
75 measurements through the differential optical absorption spectroscopy (DOAS) technique and
76 then converted to vertical columns using Air Mass Factors (AMFs). The 310.5-340 nm range in
77 OMI's UV2 channel is used in retrieving SO_2 , with focus on the 310.8 and 313 nm wavelengths.
78 The band residual algorithm (Krotkov et al., 2006) and the Linear Fit (LF) algorithm (Yang et
79 al., 2007) were first used as the OMI operational algorithms for retrieving planetary boundary
80 layer (PBL) SO_2 and volcanic SO_2 vertical column densities (VCDs) respectively. These were
81 replaced with the principal component analysis (PCA) based algorithm (Li et al., 2013) which
82 retrieves SO_2 amounts directly from spectral radiance measurements. The same technique was
83 also applied to OMI volcanic SO_2 retrievals (Li et al., 2017). This data-driven approach does not
84 rely on extensive radiative transfer modeling and has led to reduced biases and significant
85 improvements (Fioletov et al., 2015). For volcanic retrievals, algorithms still have uncertainties
86 in SO_2 mass in volcanic plumes, especially in the presence of relatively larger errors in the
87 assumed *a priori* profiles.

88 In addition to column amounts, backscattered radiances can provide important
89 information about the height of an SO_2 layer. Conceptually, a change in altitude of an SO_2 plume
90 alters the number of backscattered photons going through the layer. If a plume is high in the
91 atmosphere, more photons that are scattered from below the layer pass through the absorbing
92 SO_2 plume. This results in larger SO_2 absorption structures in the measured radiance spectra,
93 especially in the 310-320 nm range where Rayleigh scattering is dominant. Relative to the SO_2
94 amount, obtaining a fast retrieval of the height of a volcanic plume presents a greater challenge.
95 Until recently, retrieval techniques have involved a direct spectral fitting approach that use
96 backscattered ultraviolet (BUV) measurements in conjunction with extensive forward radiative
97 transfer modeling. For instance, the Iterative Spectral Fitting (ISF) algorithm (Yang et al., 2009)
98 for OMI was utilized to determine the altitude of SO_2 layer by adjusting the height while

99 minimizing the differences between measured radiances and forward RT calculations. Another
100 study has utilized an optimal estimation algorithm along with the VLIDORT radiative transfer
101 (RT) model to retrieve SO₂ density and plume height from the GOME-2 instrument (Nowlan et
102 al., 2011). Sulfur dioxide amounts and plume heights have also been estimated with the infrared
103 atmospheric sounding interferometer (IASI), through brightness temperature changes and
104 relative intensities of absorption lines (Clarisse et al., 2008). For these techniques, extensive
105 radiative transfer modeling is needed, in addition to a variety of assumptions including a
106 reasonable first guess for the plume altitude. Newer schemes were later developed for GOME-2
107 using the SOPHRI algorithm (Rix et al., 2012), a DOAS based technique that included
108 minimizing differences between plume height from simulated spectra and the assumed height
109 from measured spectra. This technique allowed for reasonably fast retrievals that could be used
110 in near real-time, thanks to the use of pre-calculated GOME spectra stored in a look up table
111 classified according to SO₂ column, SO₂ heights and other physical parameters. An updated
112 algorithm was also developed for IASI (Clarisse et al., 2014), this time implementing an optimal
113 estimation fit approach with pre-calculated Jacobians. Faster and more efficient methods for
114 GOME-2 (Efremenko et al., 2017) and TROPOMI (Hedelt et al., 2019) have made use of
115 machine learning algorithms, specifically neural networks (NNs), to develop a trained, full-
116 physics inverse learning machine (FP_ILM) for retrieving SO₂ plume height. This approach has
117 shown good accuracy and speed fast enough for near-real-time operations. The FP_ILM has also
118 been used for retrieving ozone profile shapes (Xu et al., 2017) and geometry-dependent
119 Lambertian equivalent reflectivity (Loyola et al., 2020). The primary advantage of this approach
120 is the execution speed. By separating the training phase, which involves large amounts of time
121 consuming radiative transfer computations and machine learning model training, from the
122 application phase, the desired parameter can be retrieved within milliseconds for a single satellite
123 ground pixel using the inverse model. However, similar methods of retrieving SO₂ layer height
124 have not yet been implemented for OMI. Now in this study, the FP_ILM has been applied to
125 OMI to estimate SO₂ layer height from backscattered Earthshine radiance measurements. The
126 retrieval was tested on four past volcanic eruption cases and performance was assessed through
127 machine learning metrics, as well as comparisons to other datasets such as those from
128 TROPOMI, IASI and CALIOP lidar instruments.

129

130 **2 Methodology:**

131
132 The FP_ILM approach consists of two parts, the training phase and the application (or
133 operational) phase. The training phase starts with the generation of a synthetic training dataset of
134 top of the atmosphere (TOA) reflectance spectra from a radiative transfer model. This spectral
135 dataset is then used to train a Multi-Layer Perceptron Regression (MLPR) NN model to predict
136 the SO₂ layer height as an output. In the application phase, the trained inverse model is applied to
137 real OMI radiance measurements. This inverse model is optimized from the training, and the
138 predictions of SO₂ layer height based on the model are very fast as compared with the time-
139 consuming RT calculations during the training phase. The main steps of the algorithm are shown
140 in a flowchart (Figure 1) and discussed in detail in the next sections.

141 142 **2.1 Forward Radiative Transfer Model**

143
144 The first step in the training phase is to build a large data set of synthetic backscattered
145 Earthshine reflectance spectra from forward radiative transfer (RT) calculations. These
146 calculations are performed using the Linearized Discrete Ordinate Radiative Transfer (LIDORT)
147 model with the rotational Raman scattering (RRS) capability (Spurr et al., 2008). This version of
148 the model treats first-order inelastic Raman scattering in addition to all orders of elastic
149 (Rayleigh) scattering processes. Rotational Raman scattering occurs when a photon is scattered
150 at lower or higher energy levels than the incident radiation. RRS cannot be neglected; it is known
151 to be responsible for the Ring effect (Grainger and Ring 1962), a spectral interference signature
152 characterized by the filling-in of Fraunhofer lines and telluric-absorber features. Allowing for
153 RRS in the RT model leads to differences in calculated radiances compared to those made with
154 purely elastic scattering, as characterized by the filling-in factor. This quantity is generally of the
155 order of a few percent, consistent with estimates that 4% of the total scattering in the
156 atmospheric is inelastic (Young, 1981). Fundamentally the SO₂ layer height information can be
157 retrieved by backscattered radiance spectra because the amount of scattering occurring in the
158 overlying atmosphere is determined by the height of the volcanic SO₂ plume. This is
159 demonstrated by comparing two otherwise identical RT calculations with different SO₂ layer
160 heights (Figure 2a). At shorter wavelengths where Rayleigh scattering is stronger, there is less
161 backscattered radiance for the case with higher SO₂ plume height, particularly at shorter
162 wavelengths < 320 nm (Figure 2b). Likewise, the filling-in factor (Figure 2c) shows the

163 importance of including RRS in the RT calculations as in some cases there can be 2-3%
164 difference between the Raman and elastic calculations.

165 All LIDORT-RRS calculations in this study were performed for the 310-330 nm spectral
166 range, which captures strong SO₂ and ozone absorption features. The model is supplied with
167 ozone (Daumont et al., 1992) and SO₂ absorption (Bogumil et al., 2003) cross sections,
168 atmospheric profile, ozone profile and a high resolution Fraunhofer solar irradiance spectrum.
169 The atmospheric profile has 48 layers and contains a temperature/pressure/height grid from the
170 standard US atmosphere, with an increased vertical resolution of 0.5 km below 12 km. The
171 ozone profile is determined by the total column amount, latitude zone and month as specified in
172 the TOMS V7 ozone profile climatology (Bhartia, 2002), while the SO₂ profile is assumed to be
173 a Gaussian shape with a full width half maximum (FWHM) of 2.5 km. The solar spectrum is a
174 re-gridded version of the high resolution synthetic solar reference spectrum (Chance and Kurucz,
175 2010), originally with a spectral resolution of 0.01 nm. The re-gridded version has a resolution of
176 0.05 nm, finer than that for OMI (0.16 nm sampling for a FWHM spectral resolution of ~0.5
177 nm). The advantage of using this reference spectrum over the instrument-measured irradiance is
178 that only one set of calculations is needed; they can be applied to multiple instruments and
179 instrument cross track positions without utilizing unique measured solar flux spectra for each
180 situation. Using instrument-measured solar flux data may carry less potential error and be able to
181 better handle issues with instrument degradation. However, the downside is that the inverse
182 model would need to be re-trained whenever a new measured solar flux spectrum is used. Since
183 we expect the retrieval to be primarily sensitive to SO₂ absorption signatures, the radiative
184 transfer calculation was performed for a molecular atmosphere with no aerosol scattering.

185 In order to obtain a large number of different spectra, eight key physical parameters were
186 varied for the LRRS calculations. These parameters include solar zenith angle (SZA), relative
187 azimuth angle (RAA), viewing zenith angle (VZA), surface albedo, surface pressure, O₃ column
188 amount, SO₂ column amount and SO₂ layer height. The ranges of these parameters are given in
189 Table 1.

190 The number of calculations and the parameter sets for each simulation were determined through
191 a smart sampling technique (Loyola et al. 2016). A selective parameter grid with sets of
192 parameters for each simulation was established through the use of Halton sequences (Halton,
193 1962) in 8 dimensions. The calculations are continued until the moments of the output data,

194 mean and median converged across all wavelengths. In total around 200,000 calculations were
195 done to achieve sufficiently comprehensive sample size for the variation in the eight parameters
196 across all rows of OMI. This sampling was done in order to ensure that 1) each set of parameters
197 was unique and training data is diverse; and 2) that the sample size of the entire dataset is large
198 enough for the machine learning application.

199 200 **2.2 Data pre-processing**

201
202 After the RT calculations are completed, the spectra are convolved with OMI instrument slit
203 function. Since each cross-track position of OMI contains a unique slit function, the appropriate
204 function was applied based on the VZA input for that particular calculation. The VZA ranges
205 from 0-70° across all rows in the OMI swath, with the middle (nadir) rows having a VZA of
206 close to 0. For each row, only spectra within +/- 3° of the actual VZA were convolved with the
207 appropriate slit functions. In addition, Gaussian noise with a signal-to-noise ratio (SNR) of 1000
208 was added to the spectra. While the SNR of OMI tends to be lower (Schenkeveld et al., 2017),
209 adding too much noise can greatly decrease performance of the machine learning (Table 2). The
210 root mean squared error (RMSE) and mean absolute difference (MAE) between the SO₂ height
211 from the RT calculation parameter sets and the height predicted by the neural network were used
212 as metrics (see Section 3). At SNRs of less than 500, the performance starts to increasingly
213 degrade. Between 1000 and 500 SNR, there is an increase of around 0.1 km in RMSE. However,
214 adding some degree of noise is necessary to account for errors in satellite instrument
215 measurements.

216 Next, principal component analysis (PCA) was applied to the spectral dataset for each
217 row, in order to extract the most significant features of the spectra, and to reduce dimensionality.
218 Since each convolved sample consists of 142 wavelength points, the dimensionality of this
219 problem becomes very large. However, PCA transforms each sample to a set of weights based on
220 8 principal components (PCs). These principal components explain 99.998% of the variance in
221 the synthetic dataset (Figure 3). Including additional PCs does not add any significant value to
222 the retrieval and may even lead to overfitting. Prior to starting the machine learning process, the
223 dataset is split into a training subset (90%) and a testing subset (10%). The training subset is used
224 for the neural network learning, while the testing subset only deployed verifying the performance
225 of the network to predict the output.

226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257

2.3 Machine Learning using a Neural Network

The 8 PCs, and selected parameters including the SZA, RAA, VZA, surface pressure and surface albedo were used as input for training a MLPR, which is sometimes referred to as a deep neural network. The output layer of the NN contains the effective SO₂ layer height. Column amounts of SO₂ and O₃ were not included in the training or in the application stage because of the large dependency of column amounts on SO₂ layer height and due to biases in OMI ozone retrieval in the presence of the enhanced SO₂ plume, respectively. To improve stability, the inputs (PC weights, SZA, VZA, etc.) and output (effective SO₂ height) are scaled between -0.9 and 0.9 according to the minimum and maximum of each input variable prior to input into the NN. In a NN, the input and output layers are connected by hidden layers containing neurons (also known as nodes). Each neuron is connected to others by a series of weights, by means of which the input data is passed to the next level as a weighted sum of all inputs. Inside the neural network, the Adam optimizer with a stochastic gradient descent algorithm (Kingman et al., 2014) is used to minimize the loss function, in this case the mean squared error (MSE) between the result of each iteration and the actual SO₂ layer height used to generate the synthetic spectral sample. With each iteration, the partial derivative of the MSE with respect to each node is calculated; this is used to update the weights. The training of a NN progresses by cycling through iterations of the entire training dataset, called epochs, until the training and validation MSE is minimized and there is no improvement to be obtained from further training. Throughout the training, the NN uses 10% of the training subset for validation to assess the performance with each iteration. This validation set is different from the independent test data that was set aside from training. The “tanh” (hyperbolic tangent) activation function is applied at the hidden layers to further increase stability in the NN. Other activation functions (e.g., ReLU and PReLU) were tested, however tanh was found to produce slightly better NN performance. There is also considerable flexibility in the structure of the NN, in particular the number of hidden layers and nodes in each layer. The final configuration of the NN in this study includes 2 hidden layers with 20 and 10 nodes in the first and second layer, respectively. This was determined through testing and analyzing the errors of the NN with respect to the synthetic test data set and the quality of the retrieval results after application to satellite measurements. More complex configurations of hidden layers and number of neurons were also tested and found to have worse performance

258 when using OMI data as input. Hence the relatively simple configuration was chosen as the final
259 setup for this study.

260 In neural networks a common problem known as overfitting often occurs when the
261 machine learning model is tuned so closely to the training inputs that it does not perform well on
262 new data. During training this can be diagnosed if the validation error is much higher than the
263 training error. To reduce overfitting, L2 regularization was implemented in the training. The
264 regularization reduces the effect of small and very large weight values by penalizing the MSE
265 loss function. For this study, the training was done separately for each OMI row due to the
266 different VZAs and slit functions between rows; however, the configuration of the NN was kept
267 constant between rows. The only difference in the training is the number of training epochs
268 conducted for each row before the solution becomes optimal for that row. The number of epochs
269 varies slightly but is in the 200-300 range for all rows. The final trained version of the NN, the
270 inverse operator, contains the optimal weights needed to predict the SO₂ layer height from an
271 input of separate test data.

272 An important aspect for neural network performance is the number of training samples.
273 Aside from smart sampling, the appropriate number of samples for training can be determined by
274 comparing errors from training runs where different percentages of training samples were
275 removed (e.g. 10%, 20%, 50%) beforehand. The mean absolute error between height predicted
276 by the NN and the test set height was calculated when using different numbers of input samples.
277 With a 50% reduction in training samples, the absolute error went up by around 0.3 km. In
278 contrast, reducing the training set by 10% had little impact on the error (see Table A1). These
279 results provide confirmation that for this case the training data are adequate, and that there would
280 likely be diminishing returns in NN performance with a larger training dataset.

281 **2.4 Application to satellite measurements**

282 In the application phase of the retrieval, the inverse operator is applied to OMI radiance
283 spectra, resulting in a predicted SO₂ layer height for each ground pixel in the OMI swath. For
284 this the OMI L1B Geolocated Earthshine radiance dataset is used. Since OMI only provides
285 absolute radiances, these data were normalized with respect to the same solar flux spectrum as
286 used in the generation of the synthetic spectra. In other words, the measured input becomes the
287 fraction of backscattered radiance to the incoming solar irradiance (i.e., reflectance spectrum).
288 Prior to normalizing, the irradiance spectrum was convolved with an OMI slit function for the

289 particular OMI row and orbit. The irradiance spectrum is convolved with the appropriate OMI
290 slit function in order to have consistency in wavelength points between the measured radiances,
291 synthetic radiances and irradiance of each row. To follow the same procedure as was used in the
292 training step, the PCA operator from the training phase is applied to the OMI spectra to perform
293 the dimensionality reduction and obtain a set of PC weights for each sample. The other inputs are
294 VZA, SZA, RAA, albedo and surface pressure parameters from the OMI data files. As in the
295 training phase, all inputs are scaled to the $[-0.9, 0.9]$ range. After SO₂ heights are retrieved
296 separately for each row, one height value is given for each pixel (and spectral sample). The
297 application phase of the retrieval takes only 2-3 seconds for a given row. This short duration
298 includes the application of the training phase PCA operator to OMI measurements, the scaling of
299 inputs and the deployment of the inverse operator. The whole process is repeated for each row in
300 order to get a prediction for an entire OMI swath. For some rows the retrieval is unreliable due to
301 the row anomaly, which negatively affects the quality of the OMI L1B radiance data at all
302 wavelengths and consequently L2 retrievals.

303

304 **3 Impacts of various parameters on the performance of the trained inverse model**

305

306 From the training phase, it becomes clear that the performance of the algorithm will
307 depend on several factors. As demonstrated in Fig. 3, an important factor is the SO₂ column
308 amount. Overall, the NN makes better predictions for the test data subset for SO₂ amounts > 40
309 DU. Below 40 DU, information content on the layer height to be retrieved becomes increasingly
310 small, as evidenced by large differences between predicted heights and those in the actual test set
311 (Figure 5a). Additionally, larger SO₂ loadings result in greater sensitivity between two heights,
312 as seen by comparisons of SO₂ height Jacobians for multiple amounts (Figure 4). Quantitatively,
313 if samples with SO₂ amounts less than 40 DU are excluded, the RMSE decreases from 1.48 to
314 1.15 km (Table 3). As with other sensitivity analyses, the RMSE and MAE in Table 3 are
315 calculated between the predicted output from NN and the height from the independent test
316 dataset. We can therefore expect the retrieval to produce reasonable results for moderate to large
317 volcanic eruptions. In widely dispersed plumes where the SO₂ VCD is low or for volcanic
318 degassing events, the retrieval would be less accurate. The second major dependency is on SZA.
319 The problem here stems from the occurrence of relatively large errors in RT modeling due to
320 shallow light paths and lower OMI SNR at the higher SZAs. Reasonably accurate results are to

321 be expected only for $SZA < 75^\circ$. Figure 2b shows significant differences in predicted and actual
322 heights in spectra associated with large SZAs, after removal of low VCD samples. For the final
323 training approach, it was therefore necessary to exclude spectra with large SZAs. Dependencies
324 on other physical parameters are small when compared with these two issues discussed here,
325 although there is some evidence that high surface albedo also increases error. If we remove
326 spectra with albedo > 0.6 there is a minor improvement in RMSE from 0.93 to ~ 0.89 km.
327 However, even with strong volcanic SO_2 signals, we can realistically expect that on average the
328 absolute error to be at least 1 km, due to inherent simplifications in the neural network retrieval
329 approach. The errors in actual retrievals using OMI data are expected to be larger (see Section
330 4.4).

331 **4. OMI SO_2 Effective Layer Height Results**

332 For testing the FP_ILM retrieval on OMI data, four volcanic eruption cases with sufficiently
333 strong SO_2 signals were selected (i.e. where peak SO_2 VCDs were greater than 40 DU). Each
334 case is described in detail in the following subsections. For each case, comparisons were made to
335 other satellite-derived datasets where available, for example the CALIOP lidar onboard
336 CALIPSO, the IASI SO_2 layer height retrieval (Clarisse et al., 2014), and the GOME-2
337 (Efremenko et al., 2017) and TROPOMI retrievals (Hedelt et al., 2019). It is important to note
338 that the CALIOP lidar only indicates the height of the ash plume and not the SO_2 height.
339 Although ash and SO_2 plumes are often collocated, this is not always the case, making direct
340 comparisons difficult.
341
342

343 **4.1 Kasatochi (2008)**

344 Kasatochi is a volcano located on the Aleutian Islands of Alaska ($52.178^\circ N, 175.508^\circ W$). It
345 underwent a series of eruptions beginning late in the day on August 7th, 2008, which injected
346 great amounts of ash and SO_2 into the stratosphere. Overall the explosion released roughly 2
347 million tons of SO_2 , at the time the highest SO_2 loading since the Mt Pinatubo eruption (Yang et
348 al, 2010). SO_2 effective layer heights retrieved using the machine learning model for OMI (orbit
349 21650) on August 10th, 2008, were around 11-12 km with some portions being slightly lower
350 (Figure 6a). This is in reasonable agreement with previous SO_2 height retrievals of 9-11 km
351 which used the ISF algorithm for OMI (Yang et al., 2010), considering that the uncertainty of
352

353 both retrievals are around 2 km. Likewise, Nowlan et al. (2011) showed that the majority of the
354 plume was around 10 km, and up to 15 km in some parts. There is also agreement with IASI
355 (Figure 6b) and CALIOP data (Figure 6d) which showed plume heights of 10-12 km and 12.5
356 km respectively. It is important to note that the IASI overpass occurred later in the day than those
357 for OMI and CALIPSO. Another verification source we used was the GOME-2 SO₂ layer height
358 retrieval that uses FP_ILM (Efremenko et al., 2017). The study found a height of around 10 km
359 and up to 14 km in areas of high SO₂ loading for August 10th (Figure 6c). The GOME-2 overpass
360 occurred 4 hours earlier than OMI. The mean, median, standard deviation and the inner quartile
361 range (IQR) of the three retrievals (Table 4) also show good agreement for this case. Although
362 the OMI results agree well in general with the results of these studies and datasets, the retrieval is
363 less sensitive with respect to detecting variability in the SO₂ layer height within the plume.

364

365 **4.2 Kelud (2014)**

366 Kelud, a stratovolcano located in East Java, Indonesia (7.935°S, 112.315°E), erupted on
367 February 13th, 2014 at 1550 UTC, in the process depositing ash in a 500 km diameter around the
368 volcano and leading to mass evacuations from nearby towns. Even though this case has
369 somewhat lower SO₂ VCDs than those from Raikoke and Kasatochi, the peak SO₂ VCDs of ~60-
370 70 DU should still allow for retrievals with reasonable accuracy (see section 2). The OMI
371 retrieval results indicate that the maximum height of the main plume was 18-19 km (Figure 7a),
372 although other studies suggest that several smaller layers of SO₂ and ash were located as high as
373 26 km (Vernier et al., 2016) on the previous day. However, the SO₂ loading at that level was
374 most likely too low for an accurate retrieval using OMI radiances. CALIOP lidar detected ash
375 plumes at around 19.5 km and the IASI retrievals registered the plume at 17.5 km over the same
376 area as that for OMI. The height of the ash plume from this eruption was also estimated using
377 Multifunctional Transport Satellite (MTSAT 2) observations and transport modeling (Kristiansen
378 et al., 2015). That study found an injected height of around 17 km, which is in agreement with
379 the OMI result, especially when considering the most probable heights on the PDF (Figure 8b).
380 We note here that only a small portion of the plume was retrieved with our algorithm, given the
381 relatively low SO₂ VCDs and interference due to the OMI row anomaly. It is promising to note
382 that the OMI retrieval was able to identify heights at the upper end of the height range used in
383 the training phase. On the other hand, while the retrieval can extrapolate to heights above 20 km,

384 the accuracy would likely degrade due to the lack of training data with heights outside of this
385 limit.

386 387 **4.3 Calbuco (2015)**

388
389 The Calbuco volcano is located in Chile (41.331°S, 72.609°W). The primary eruption
390 had a volcanic explosivity index (VEI) of 4 and occurred on April 22nd with little warning. The
391 primary plume ascended higher than 15 km, while plumes from smaller subsequent eruptions
392 stayed in the troposphere. The volcanic plume spread northeast in the following days, resulting in
393 flight cancellations at Uruguayan and south Brazilian airports. The OMI-retrieved SO₂ effective
394 layer heights in the area of greatest VCD was in the 15-17 km range. In the same region, IASI
395 results (Figure 7c) show similar plume heights, approximately around 15 km, although as with
396 the previous events, the overpass times of the two instruments are different. CALIOP lidar shows
397 the ash plume to at roughly 17 km (Figure 7e). Unfortunately, the overpass of CALIPSO occurs
398 over an area of OMI's swath that is affected by the row anomaly, and this makes a direct
399 comparison unfeasible. Nevertheless, the CALIPSO aerosol layer height is still comparable to
400 OMI-retrieved effective SO₂ layer heights for the portion of the plume further to the west. The
401 retrieval for OMI is consistent with the other instruments for SO₂ plumes, with the exception of
402 that part of the plume with SO₂ below 30-40 DU (see Figure A1), for which results were not
403 plotted in Figure 7a due to lower biases.

404 405 **4.4 Raikoke (2019)**

406 The eruption of the Raikoke stratovolcano (48.2932°N, 153.254°E), located on the Kuril Islands
407 of Russia, occurred on June 21st, 2019 at 1800 UTC. A series of explosions during the eruption
408 sent large amounts of ash and SO₂ into the lower stratosphere. Maximal loadings of SO₂
409 measured by OMI and other sensors exceeded 500 DU. In the following days the plume
410 underwent dispersion and spread out over the northern Pacific Ocean and later over eastern
411 Russia. Early estimates of plume injection height for the eruption were predominantly in the 10-
412 13 km range with potentially larger heights in some areas of the plume. In Figures 9a and 9b, the
413 SO₂ effective layer heights retrieved from OMI data are shown for the Raikoke plume on June
414 23rd and June 24th respectively. The plume heights for both days are predominantly in the range
415 10-12 km, although some areas of the plume had estimated peak heights of 13-14 km. In

416 comparison, the TROPOMI results show slightly larger heights (13-14 km) for June 24th and
417 similar heights to OMI for June 23rd (Figure 9c and d). The IASI SO₂ height product also shows
418 fairly good agreement, with heights mainly at the 10-11 km level (Figure 9e and f). It is also
419 useful to look at a distribution of heights predicted for the domain (Figure 10) in order to get a
420 more quantitative comparison between the datasets. Based on this distribution, there is clearly at
421 least 2 km difference between the most probable heights from OMI and those from TROPOMI
422 for June 24th (Figure 10b and d) and slightly lower heights in the distribution for IASI. This is
423 also displayed in Table 4 which shows a 2-3 km difference in the mean and median of retrieved
424 heights between OMI and TROPOMI. Additionally the IQR and standard deviation provide a
425 quantitative measure of the variation in the distribution of the retrieved heights, which can
426 change from one orbit to another. Note that points with lower than 30 DU are not included in the
427 PDFs for all sensors. The results are also compared with CALIOP lidar onboard CALIPSO,
428 which shows ash plume heights of 12-13 km for both days (Figure 11a and 11b). Although there
429 is overestimation for some OMI pixels, especially for June 24th, the section of the plume with the
430 CALIPSO flyover has similar heights (around 12.5 km) to lidar-determined aerosol layer
431 altitudes. Lastly, we note that a recent study highlighted probabilistic height retrievals using the
432 Crosstrack Infrared Sounder (CrIS) for Raikoke. This study found a median height of 10-12 km
433 across a large part of the plume, however with some areas upwards of 15 km. While there are
434 some notable differences across all of the datasets, the OMI retrieval for this case falls within the
435 general consensus of plume height estimates for this volcanic event.

436 437 **4.5 Discussion of errors**

438 It is clear that predicting SO₂ layer height with FP_ILM is an efficient process, but one that is not
439 flawless in terms of accuracy. As comparisons between instruments/retrievals have shown, on
440 average there were 1-2 km differences in heights, especially for the Raikoke event, although we
441 consider this to be good agreement given the estimated MAE and RMSE associated with this
442 retrieval. In this regard, the retrieval is an approximate estimate of the SO₂ plume height rather
443 than a precise determination. Differences in the retrieved heights between different
444 studies/algorithms result from differences in instruments, forward model assumptions and
445 retrieval techniques as well as uncertainties in each retrieval. For instance, IASI is a thermal IR
446 instrument and its retrieval does not use FP_ILM. Therefore exact agreement with IASI results is

447 difficult to achieve, especially since the IASI retrieval itself has a stated error range of ± 2 km,
448 although its retrievals serve as a good verification dataset. The stated uncertainty for TROPOMI
449 retrievals (Hedelt et al., 2019) is ~ 2 km for SO₂ amounts of greater than 20 DU, similar to our
450 estimated uncertainties for OMI. While the general retrieval approach for TROPOMI (Hedelt et
451 al., 2019) is similar to that for OMI in the present study, there are also important instrument
452 differences that can lead to differences in the retrieved heights between the two instruments, such
453 as the pixel size, noise, radiometric accuracy and the level of degradation. TROPOMI has a
454 much finer spatial resolution compared to OMI, with footprints typically 5.5×3.5 km² up to
455 maximum size 7×3.5 km²; TROPOMI also has larger maximal SO₂ signals. Consequently,
456 TROPOMI is better able to resolve localized variations in the height throughout the plume, and
457 is likely to be more accurate overall due to better SNR. However, current TROPOMI L1 data are
458 known to have issues with instrument degradation and radiometric accuracy in the UV spectral
459 range (Ludewig et al., 2020); this could be a potential contributing factor the differences between
460 the two instruments. OMI retrievals show more or less uniform height levels across the entire
461 plume with the peak heights in areas with the best SO₂ signal. Note, CALIOP lidar profiles
462 sometimes show disagreements with OMI retrieved heights, because CALIOP only identifies the
463 height of the ash or aerosol plume. It also offers a comparison for only a single cross section of
464 the entire plume per orbit. Despite the uncertainties, the consensus provided by different
465 instrumental datasets can provide a reasonable estimate for the SO₂ layer height, and if done in
466 near real time, can aid in decision making with regards to aviation safety.

467 Another source of error is present in the training phase. One difficulty here is finding the
468 ideal choice of neural network setup. With many parameters to consider, such as the number of
469 input PCs, number of layers, number of nodes, learning rate, regularization, weight initialization,
470 etc., it is very time consuming to optimize the neural network setup. We have found a relatively
471 simply configuration that performed reasonably well with both test data and real OMI
472 measurements for all scenarios and events considered. However, even after optimization of the
473 parameters, random error inherently exists in the neural network. A measure of random error can
474 be obtained by altering the random state of the neural network whilst keeping other parameters
475 constant. For ten trial runs with different random seeds, variations of the MAE error were around
476 0.15-0.2 km (see Table A2). Although the differences in the errors calculated with the synthetic
477 test data are relatively small, larger changes can be expected during the application phase.

478 Indeed, when applying the inverse models to OMI, there is noticeable, up to 1 km variation in the
479 retrieved height for the same pixels. It is thus difficult to improve results further than ~ 1 km
480 absolute error, even in the training phase. In the application phase, some additional error comes
481 from the differences between synthetic spectra and real satellite measurements with noise errors.
482 For example, with an SNR of 500 used in training, which is a typical noise level for OMI, the
483 RMSE of the neural network prediction is around 1.25 km (Table 3). This can be considered the
484 lower limit of retrieval error when the inverse operator is used on OMI measurements. Lastly,
485 some deviations between the measured and synthetic training spectra originate from the RT
486 modeling. The calculations contain several assumptions including the SO₂ plume shape,
487 atmospheric profiles, gas profiles, and a molecular scattering atmosphere. Further testing is
488 required in order to determine if the inclusion of aerosols in RT calculations would improve the
489 algorithm performance.

490

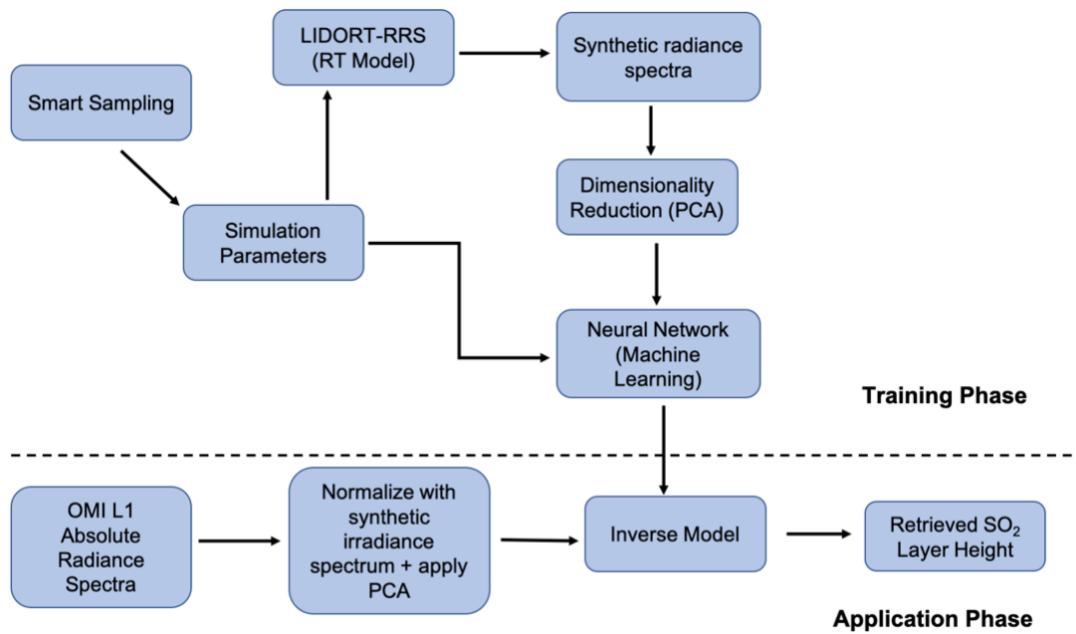
491 **5 Conclusion**

492

493 In this study we have introduced a new algorithm for OMI retrievals of the volcanic SO₂
494 effective layer height from UV earthshine radiances. This algorithm is based on an existing
495 FP_ILM method which combines a computationally time-consuming training phase with full
496 radiative transfer model simulations and a machine learning approach to develop a fast inverse
497 model for the extraction of plume height information from radiance spectra. Fast performance
498 means that the algorithm can be considered for operational deployment, given that the retrieval
499 of a SO₂ layer height prediction from the inverse model takes only a matter of milliseconds for a
500 single OMI ground pixel. For the training, a synthetic dataset of earthshine radiance spectra were
501 created with the LIDORT-RRS RT model for a variety of conditions based on choices of 8
502 physical parameters determined with smart sampling techniques. A dimensionality reduction was
503 performed through PCA in order to reduce the complexity of the problem and to separate those
504 features that best capture the great majority of variance of the dataset; 8 principal components
505 were sufficient for this purpose. Dimensionally-reduced data together with the associated
506 parameters were used to train a double hidden-layer neural network to predict SO₂ plume height
507 from any given input data. The PCA from the training phase and the inverse operator resulting
508 from the optimal NN framework were then applied to real satellite radiance spectra and
509 parameters to get retrieved values of SO₂ plume heights for several volcanic eruption events.

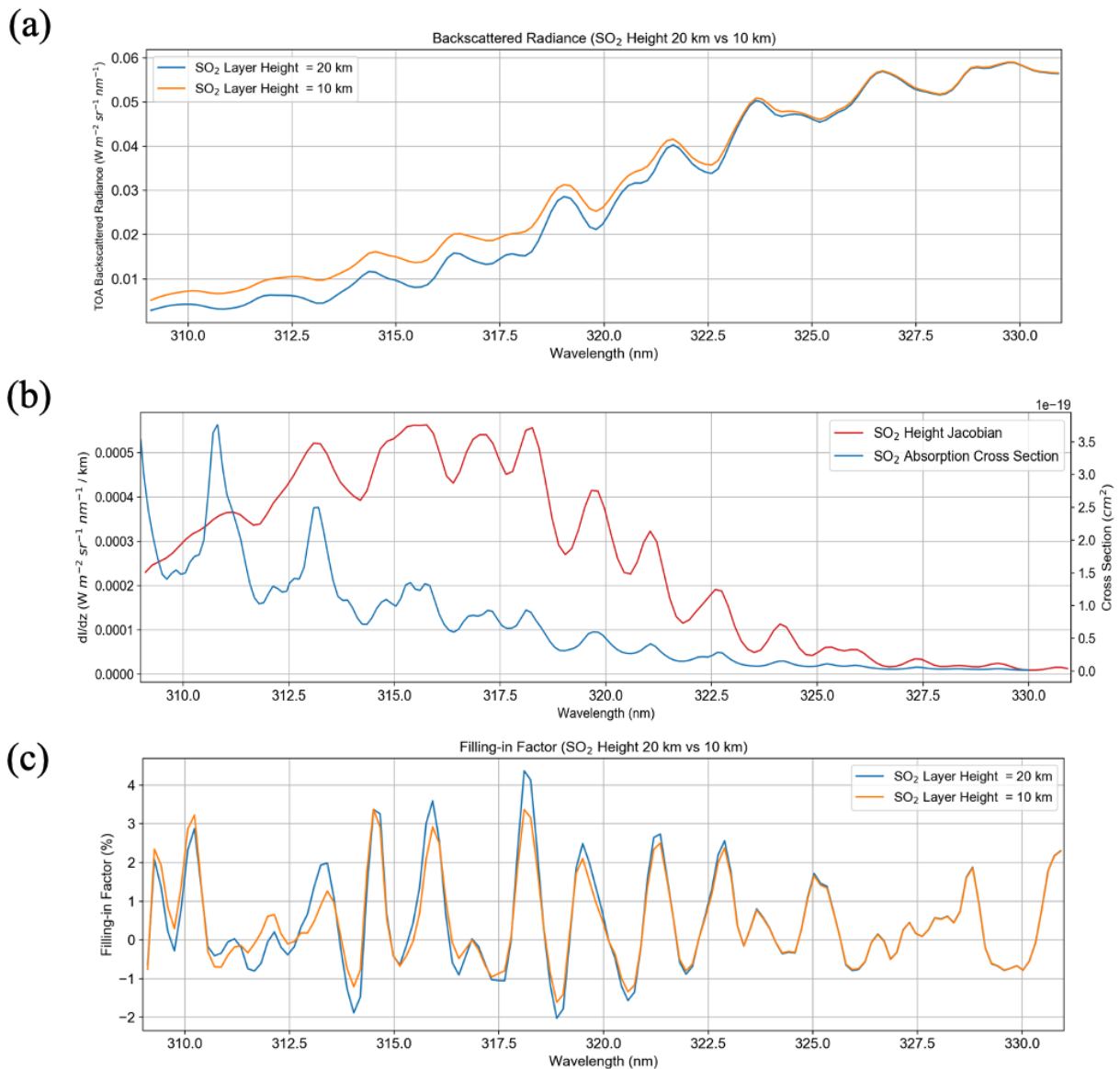
510 Through comparisons with CALIPSO lidar overpasses, as well as TROPOMI and IASI
511 retrievals, it was shown that the retrieval for OMI can estimate reasonable SO₂ layer height for
512 all the events considered, with absolute errors in the range of 1-2 km. These results can give an
513 indication of plume heights achieved during medium- to large-scale eruptions, and guide
514 important decisions in aviation hazard mitigation. For all events treated in this study, there was
515 general agreement with CALIOP lidar, although SO₂ could not be retrieved for the locations of
516 the CALIPSO flight path for the Kelud and Calbuco cases due to OMI row anomaly issues.

517 Uncertainties and sources of error in using this approach open up possibilities for future
518 work in improving the accuracy of the retrieval. We assumed that ash and sulfur dioxide plumes
519 are mostly collocated when using CALIPSO as a source to verify the plume height. Although
520 this is often true, dispersion of the plume in the days following the eruption can separate the two
521 components. Therefore, tracking these plumes become challenging when using reflectance
522 spectra alone; further analysis may need to include trajectories or wind data. The model was
523 trained on synthetic spectra calculated for molecular atmosphere conditions in the absence of any
524 aerosol loading. The impact of including aerosols in the simulations is another subject for a
525 follow-up study. We also intend to generate data sets of synthetic spectra by using a vector RRS
526 model to account for polarization. For improving the performance and efficiency of the machine
527 learning, the use of neural network ensembles and a further optimized setup of NN structure and
528 parameters will be explored. Other future work will include extending the application of FP_ILM
529 to the Suomi-NPP OMPS instrument as well as exploring the ability to predict multiple outputs
530 simultaneously from this approach.

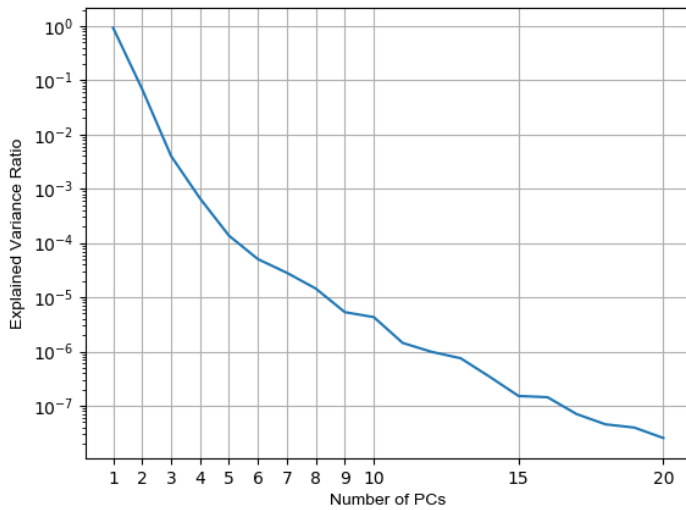


531

532 **Figure 1:** The flowchart of the FP_ILM methodology for retrieving OMI SO₂ Effective Layer Height.
 533 The steps above the dashed line are part of the training phase done prior to incorporation of OMI
 534 measurements. The application phase involves deployment of the trained model to the OMI radiance
 535 measurements to obtain estimates of effective volcanic SO₂ layer heights.
 536

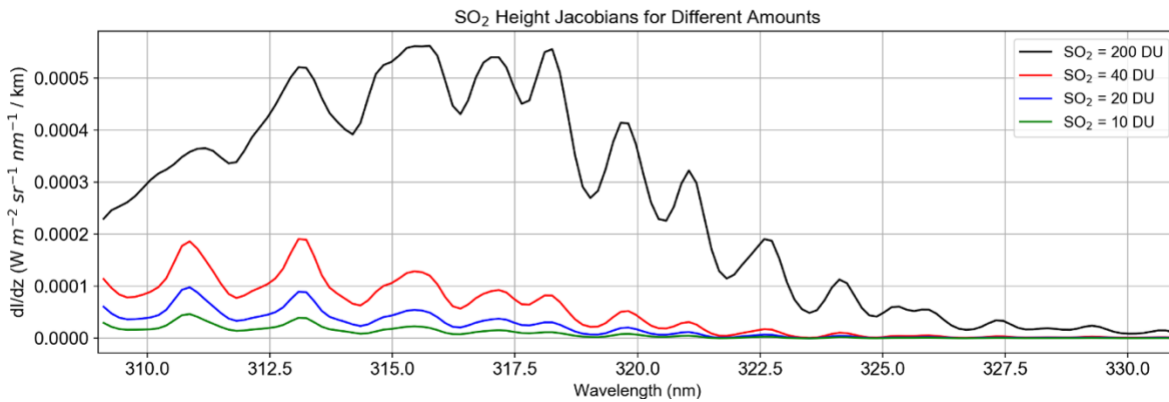


537
 538 **Figure 2:** (a) Simulated top of the atmosphere (TOA) Earthshine radiances for two different SO₂ layer
 539 heights (10 km and 20 km) from the LIDORT-RRS model. Also shown: (b) the SO₂ height Jacobian
 540 (change in radiance per km between the two spectra) along with the absorption cross-sections of SO₂ for
 541 reference; (c) the filling-in factor. The filling-in factor is defined as the difference between the total and
 542 elastic-only radiance results, divided by the total radiance, expressed as a percentage. An SO₂ column
 543 amount of 200 DU was used in the two calculations and all other parameters were kept constant except
 544 for the SO₂ layer height.
 545
 546



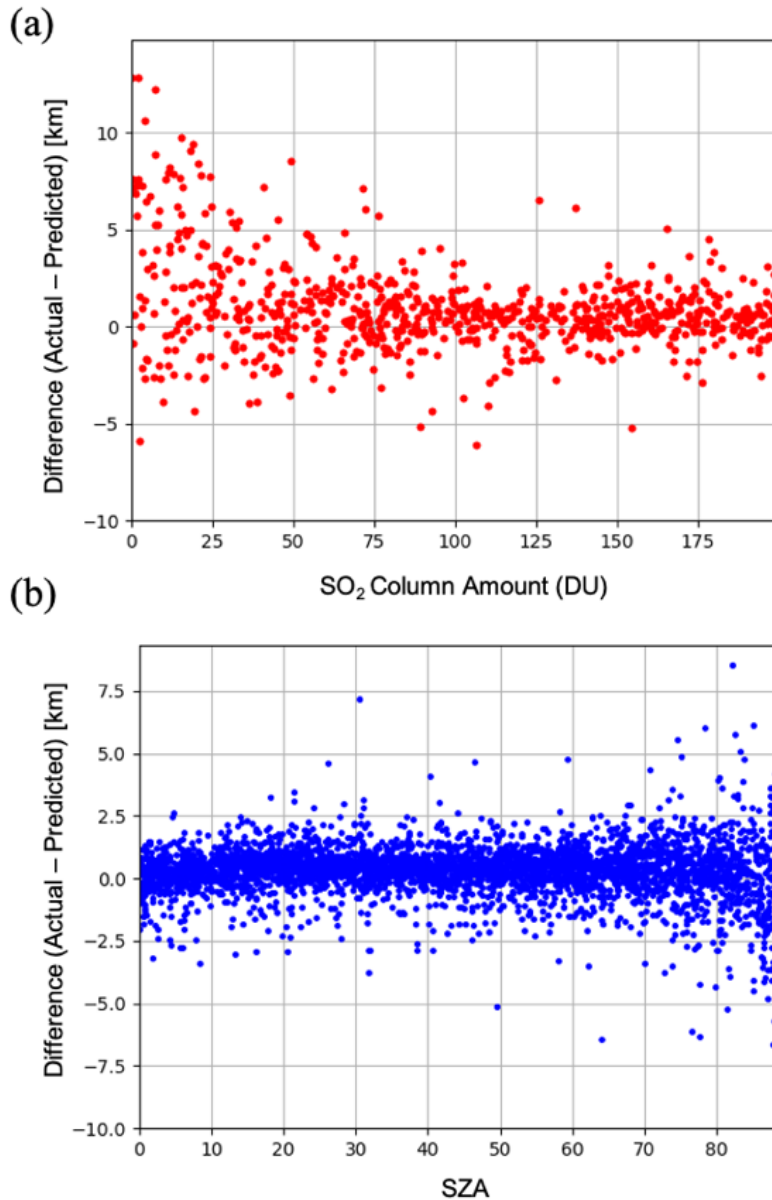
547
548
549
550

Figure 3: Explained variance ratio as a function of the number of principal components of the spectral dataset.

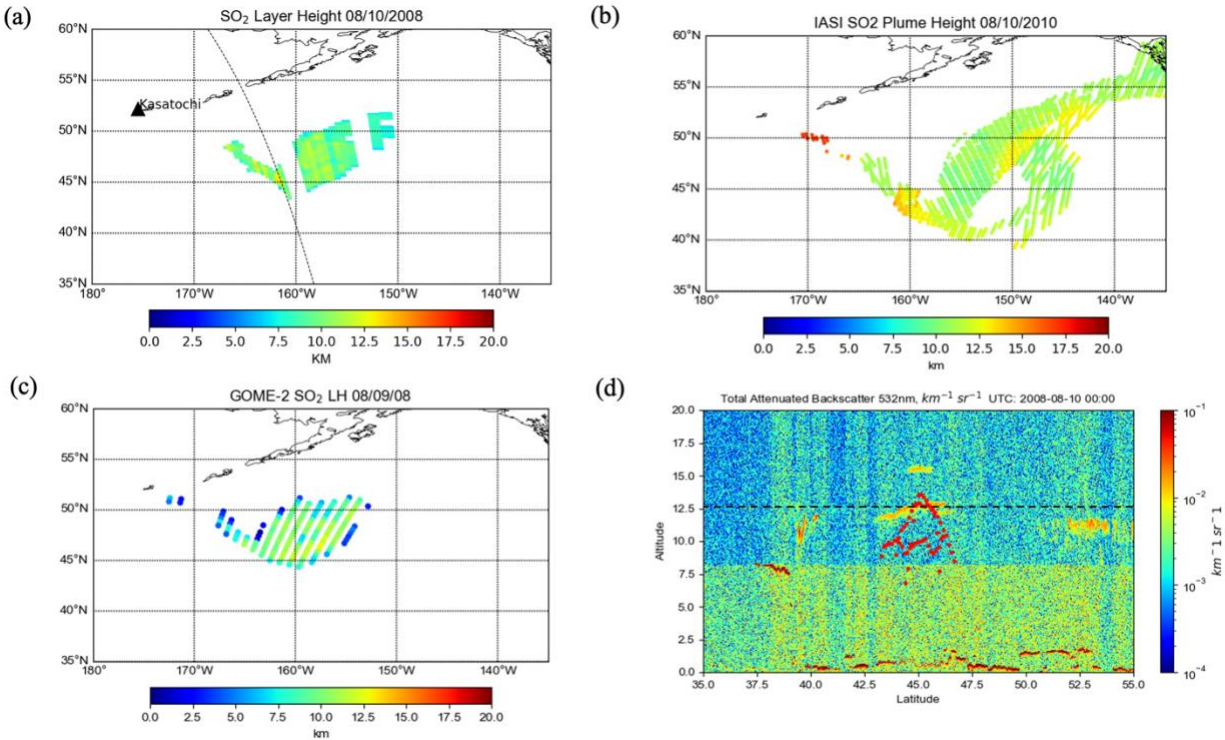


551
552
553
554
555
556

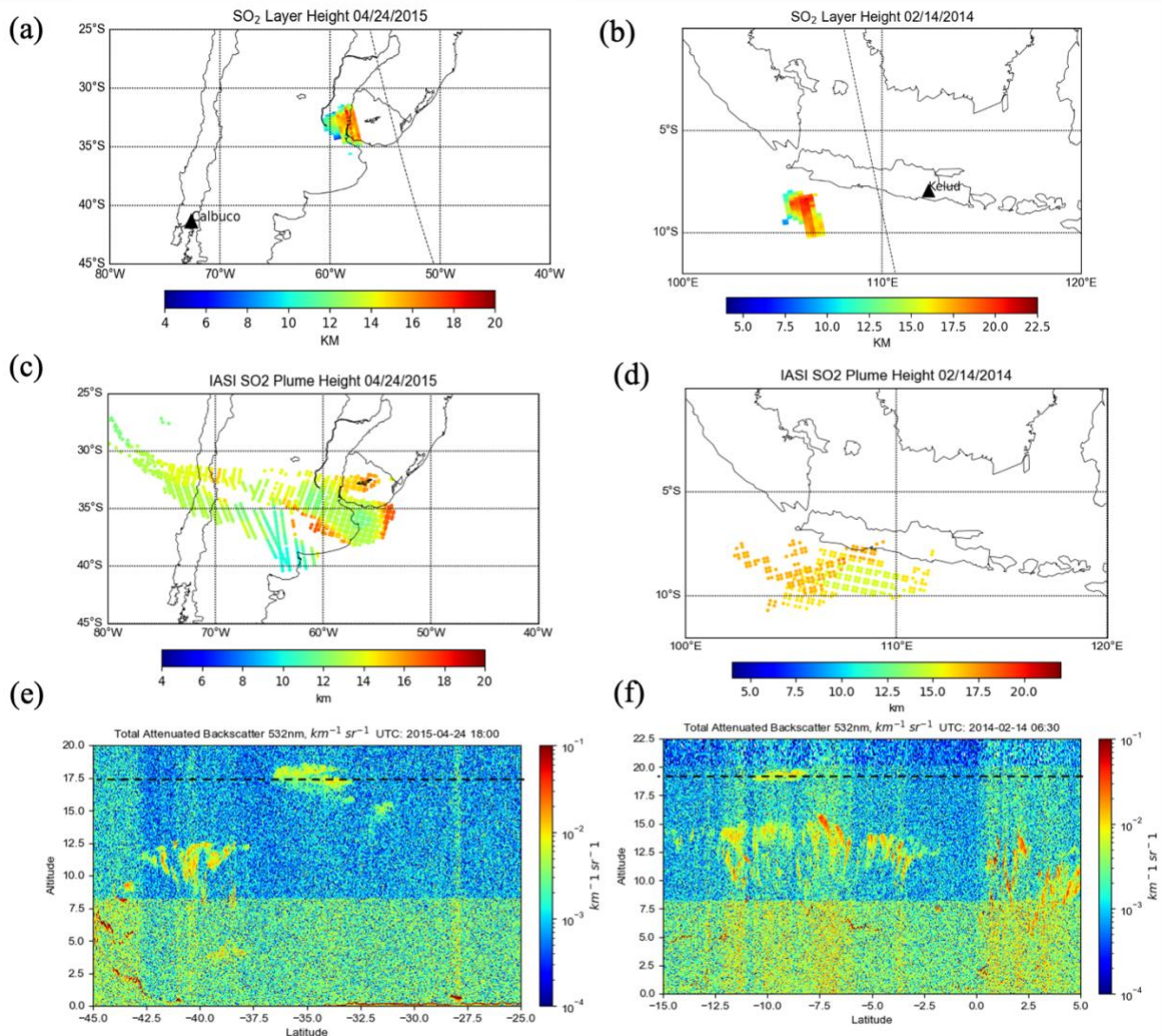
Figure 4: SO₂ Height Jacobians (dI/dz) for 4 different assumed SO₂ column amounts. The Jacobians were calculated from the difference between two radiance spectra with 10 km and 20 km SO₂ height. All other physical parameters were identical in the calculation of the spectra.



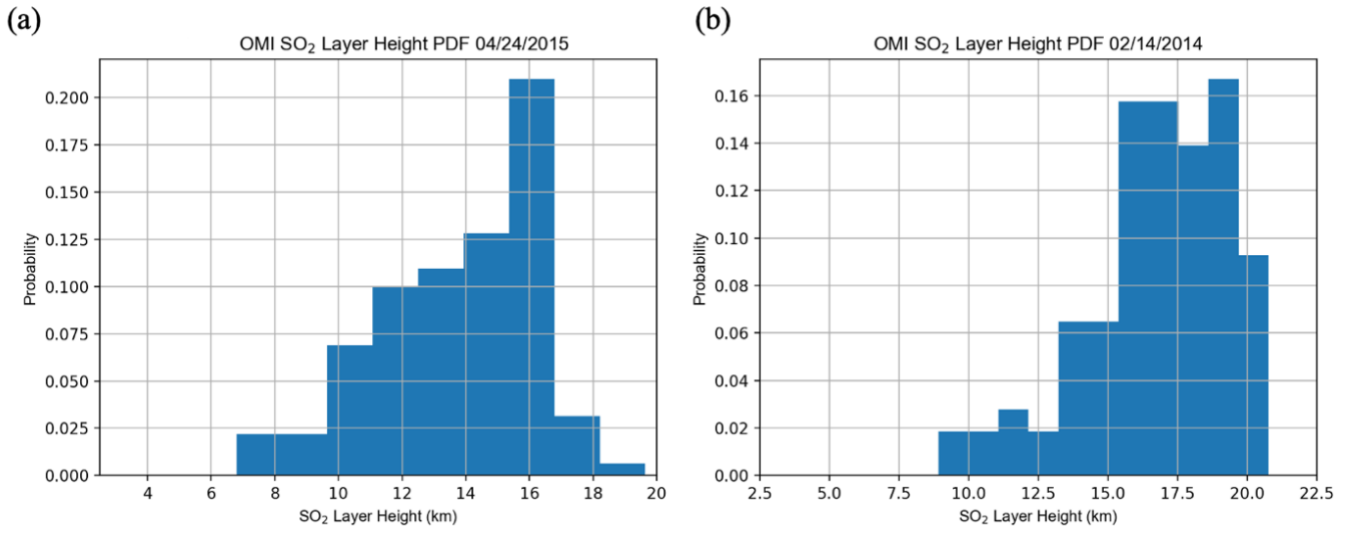
557
 558 **Figure 5:** Dependence of retrieval errors on (a) SO₂ amount and (b) SZA for cases with SO₂ VCD > 40
 559 DU. The error is defined as the difference between the SO₂ layer height predicted by the neural network
 560 using inputs from the independent test set, and the actual height from the same samples. The test set
 561 comprises 10% of the original spectral dataset withheld from training the neural network. The plots show
 562 that the retrieval error is mostly within +/- 2.5 km for SZA < 70, but increases significantly for large
 563 SZAs.
 564



565
 566 **Figure 6:** Comparison between the volcanic plume heights from (a) OMI, (b) IASI, (c) GOME-2
 567 and (d) CALIOP lidar 532-nm attenuated backscatter, for the 2008 Kasatochi eruption. The black
 568 dotted line in (a) shows the CALIPSO track. Some rows of OMI in this case were affected by the
 569 row anomaly, as seen by the gaps in the plume. The red dots in (d) show the OMI retrieval near
 570 the CALIPSO path and the black dashed line denotes the height of the ash plume observed by
 571 CALIPSO.
 572

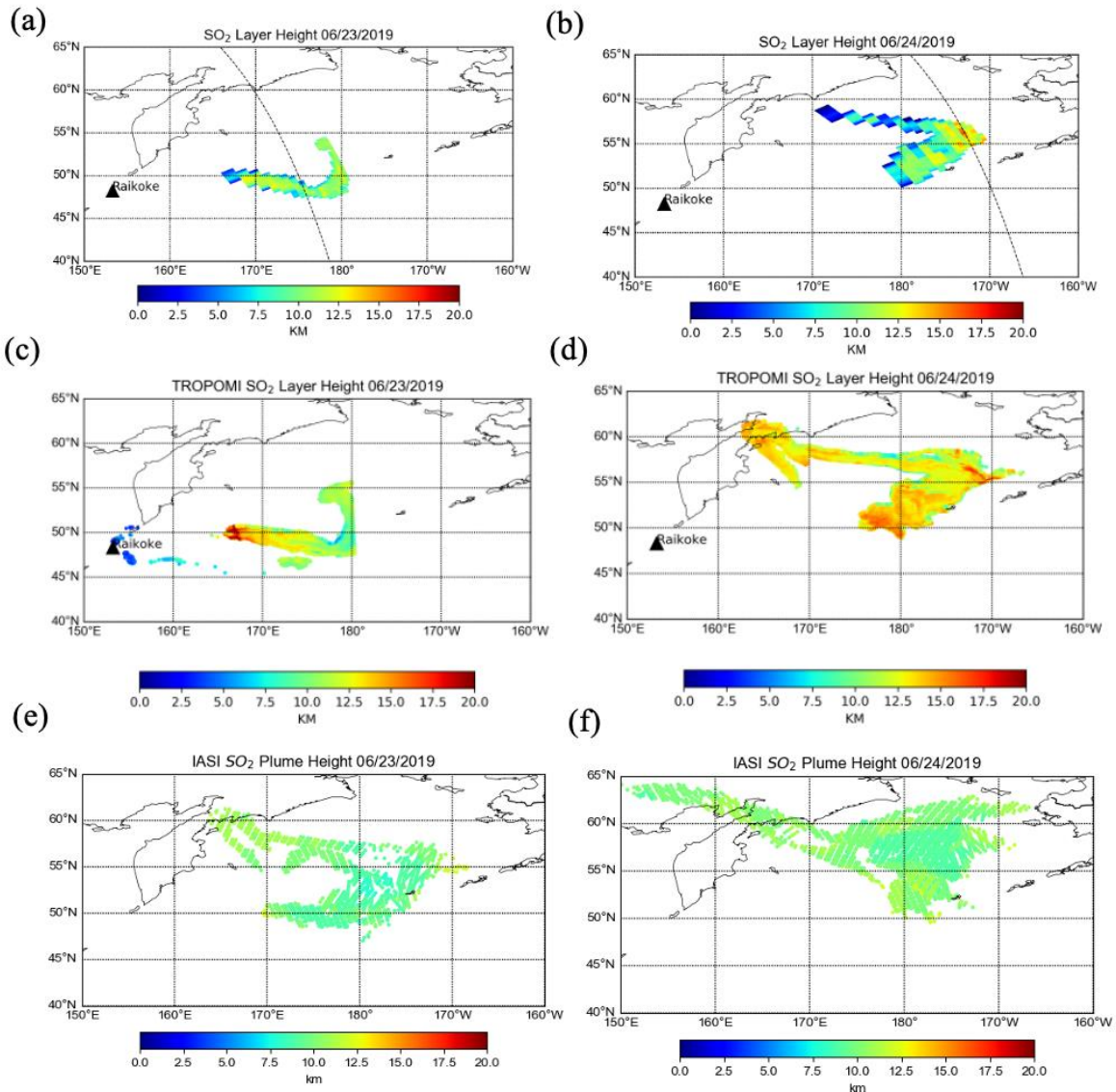


573
 574 **Figure 7:** Comparisons of plume heights for the 2015 Calbuco eruption (left) and the Kelud
 575 eruption (right) for OMI (a,b), IASI (c,d) and 532-nm total attenuated backscatter from the
 576 CALIOP lidar (e,f). For OMI, only pixels with > 30 DU of SO₂ are shown and retrievals were
 577 unavailable for some parts of the plume due to the row anomaly. The black dotted line in (a) and
 578 (b) marks the CALIPSO track. The white rectangles in (e) and (f) show the location of the plume
 579 in the lidar profile. Direct comparison with CALIPSO was not possible due to obstruction by the
 580 OMI row anomaly
 581

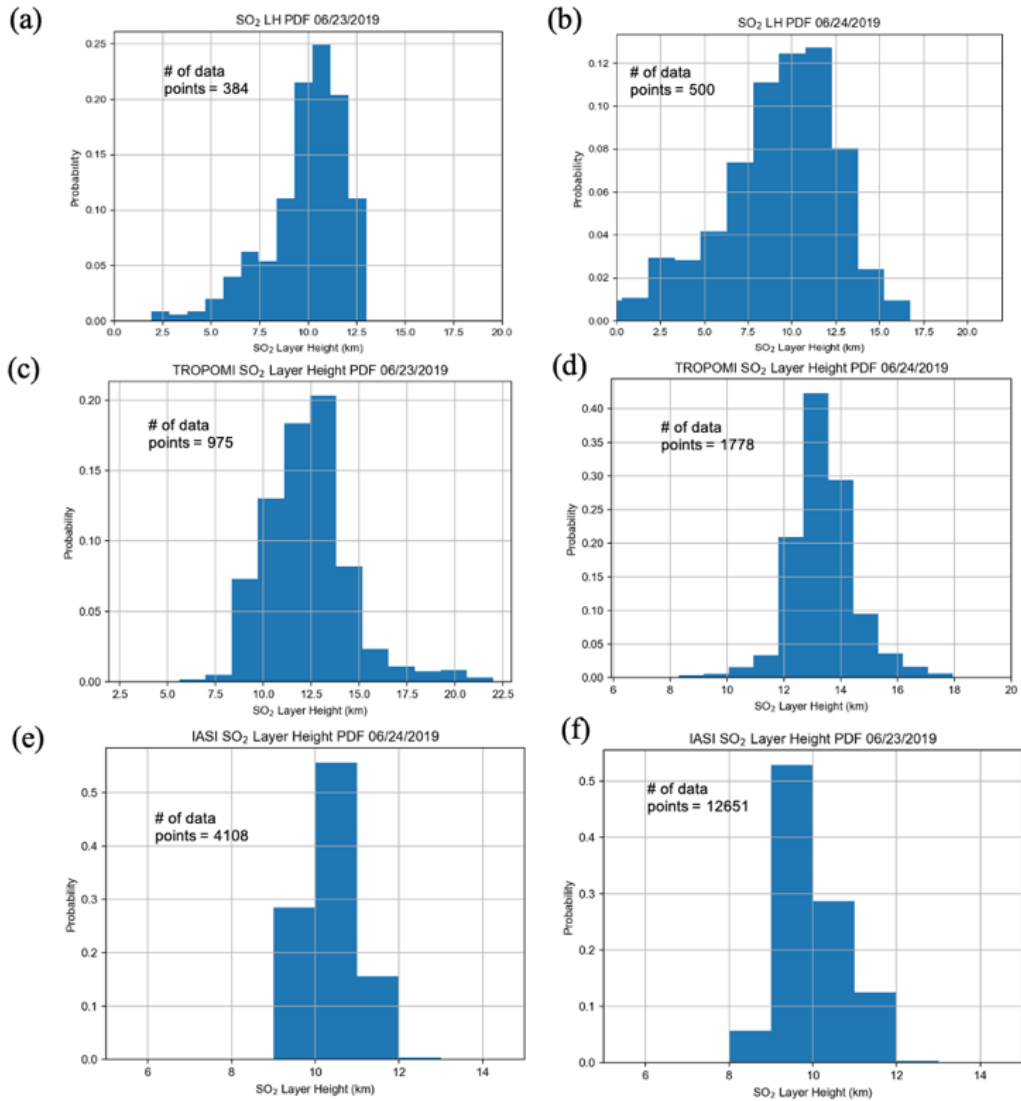


582
 583
 584
 585

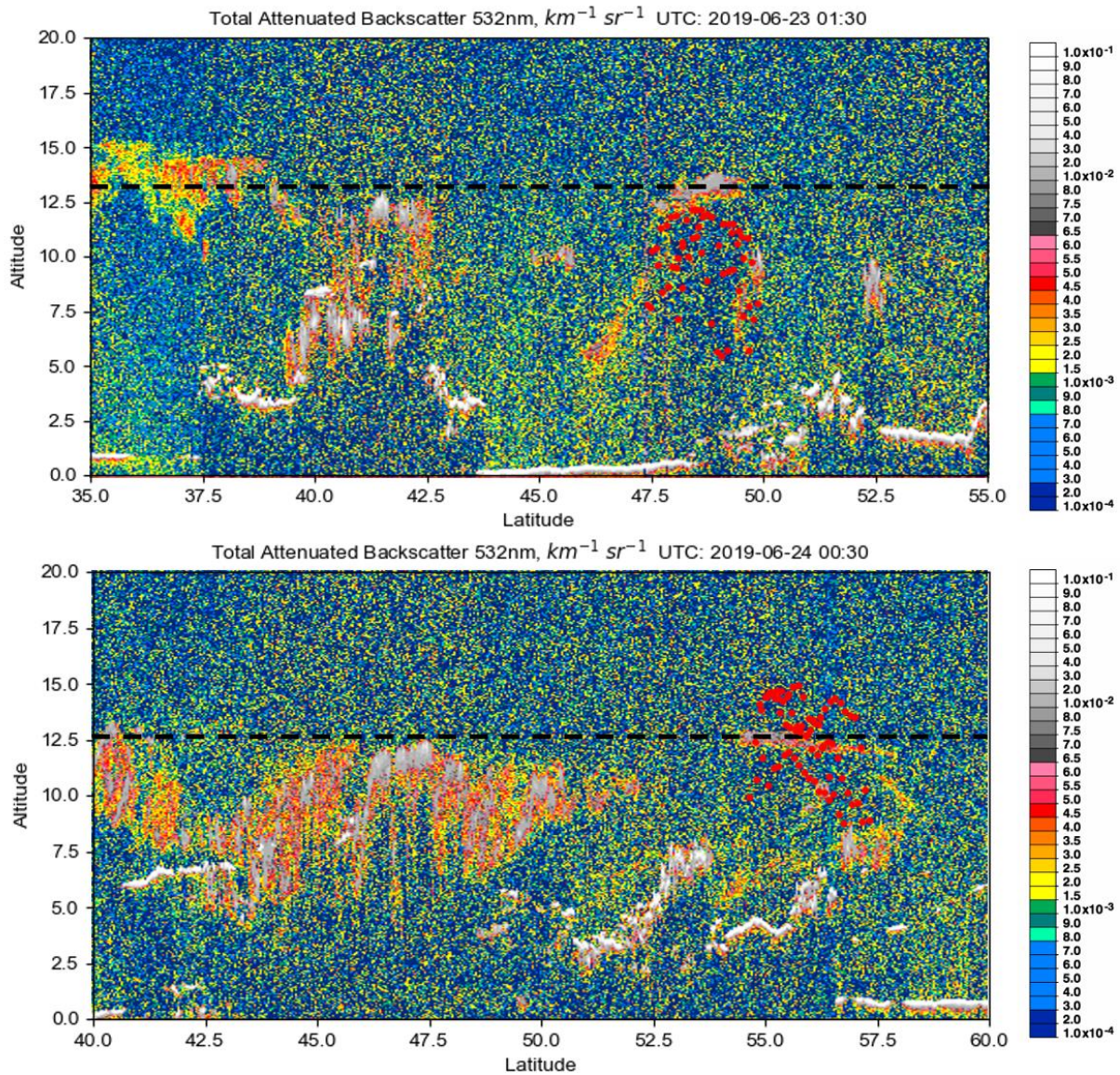
Figure 8: Probability histograms of SO₂ effective layer height retrievals for (a) the Calbuco eruption on April 24, 2015 and (b) the Kelud eruption on February 14, 2014.



586
 587 **Figure 9:** The SO₂ layer height retrieval for the Raikoke eruption plume on June 23rd, 2019 (left)
 588 and June 24th, 2019 (right) for the OMI (a, b), TROPOMI (c, d) and IASI (e, f) instruments. For
 589 all 3 sensors, only pixels where SO₂ VCD > 30 DU are shown.
 590



591
 592 **Figure 10:** Probability histograms of SO₂ layer height retrievals for (a,b) OMI and (c,d),
 593 TROPOMI on June 23rd, 2019 (left) and June 24th, 2019 (right) and (e,f) IASI. Only pixels with
 594 SO₂ column amount greater than 30 DU are included. These plots correspond to the results
 595 plotted in Figures 4a-f.
 596



597
 598 **Figure 11:** CALIPSO lidar 532-nm attenuated backscatter for the Raikoke eruption on (a) June
 599 23rd and (b) June 24th, 2019. The black dashed line symbolizes the height of ash plume seen by
 600 CALIPSO and red dots show the results from the OMI retrieval along CALIPSO's flight path.
 601 The flyovers occurred shortly after 01:30 and 00:30 UTC on June 23rd and 24th respectively,
 602 around the same time as OMI.

603
 604
 605
 606
 607
 608
 609
 610

611 **Table 1:** Ranges of the eight physical parameters varied in LIDORT-RRS for the synthetic
 612 spectra calculations.

Parameter	Range
Solar Zenith Angle	0-90°
Viewing Zenith Angle	0-70°
Relative Azimuth Angle	0-180°
Surface albedo	0-1
Surface pressure	250-1013.25 hPa
O ₃ VCD	225-525 DU
SO ₂ VCD	0-1000 DU
SO ₂ Layer Height	2.5-20 km

613

614 **Table 2:** The RMSE and the mean absolute difference (km) of all data points in the independent
 615 test set after adding noise as indicated by different SNR values. All other parameters and input
 616 data were kept constant. SZA < 75 degrees and SO₂ VCD > 40 DU were excluded from the test
 617 set for these comparisons.

	No noise	SNR=1000	750	500	200	100
Mean Absolute Difference (y_known - y_pred) (km)	0.894	0.904	0.939	0.996	1.114	1.362
RMSE (km)	1.454	1.498	1.521	1.632	1.807	2.143
R-coefficient	0.988	0.985	0.983	0.980	0.972	0.955

618

619

620

621 **Table 3:** The RMSE and the mean absolute difference of all data points in the test set under
 622 different conditions. For each condition, the appropriate points were removed and excluded in
 623 error calculations. All cases in this table used synthetic training spectra with added SNR 1000.

	All cases	SO ₂ > 20 DU	SO ₂ > 40 DU	SO ₂ > 60 DU	SZA < 75°	SO ₂ > 40 DU and SZA < 75°	Albedo < 0.6	SO ₂ > 40 DU , SZA < 75° , Albedo < 0.6
RMSE	1.487	1.216	1.150	1.109	1.281	0.931	1.524	0.895
Absolute Mean Difference (km) (Predicted – Actual)	0.910	0.834	0.803	0.782	0.795	0.697	0.895	0.667

624

625 **Table 4:** Statistical comparisons of the SO₂ height retrievals for two days of the Raikoke
 626 eruption and the Kasatochi eruption cases.

Metric (km)	Raikoke (June 23 rd , 2019)			Raikoke (June 24 th , 2019)			Kasatochi		
	OMI	IASI	TROPOMI	OMI	IASI	TROPOMI	OMI	IASI	GOME-2
Std. Deviation	1.67	0.85	1.96	2.38	0.65	1.04	1.39	0.72	1.29
Median	10.60	9.00	12.10	10.30	10.00	13.24	9.70	10.00	10.21
Mean	10.20	9.63	12.15	10.00	9.83	13.30	9.84	10.40	10.02
IQR	1.79	1.00	2.71	2.68	1.00	1.20	1.36	1.00	1.67

627
628

629 **Appendix A**

630

631 **Table A1:** Mean absolute difference and RMSE for different reductions of the original training
 632 dataset. The test was performed on training sets for five different OMI rows and the errors were
 633 averaged.

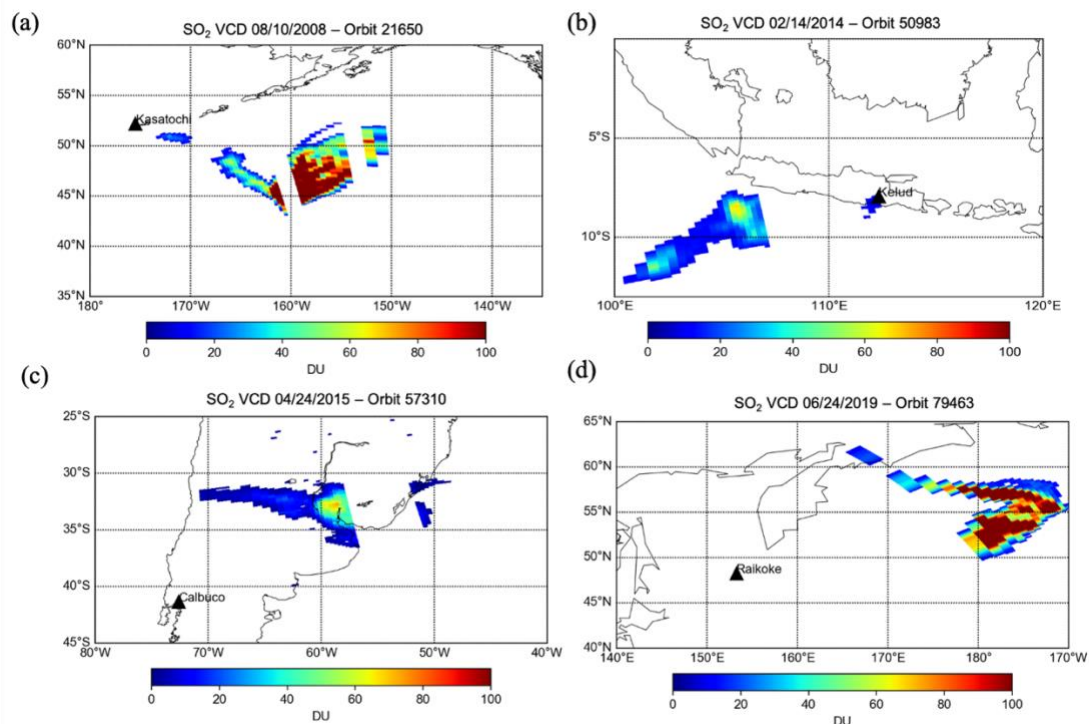
% of samples withheld	0	10	20	30	40	50
Mean Abs Difference	0.95	0.98	1.02	1.08	1.12	1.24
RMSE	1.46	1.45	1.62	1.69	1.79	2.00

634
635

636 **Table A2:** Effect of altering random seed number on error obtained using the test dataset, and
 637 the SO₂ height retrieval result after application to OMI. For the results, heights for two different
 638 pixels within the orbit from the Raikoke event (June 24th, 2019) are shown. Heights were
 639 retrieved using separate inverse models trained using 10 random states.

		Random Seed Number	1	2	3	4	5	6	7	8	9	10
NN Training error	Abs. Mean Error		0.98	1.14	1.03	1.16	1.08	1.18	1.05	1.01	1.12	0.98
	RMSE		1.69	1.85	1.71	1.78	1.79	1.92	1.71	1.67	1.73	1.70
Application (Raikoke - OMI Orbit 79463)	Sample pixel 1		10.52	10.69	10.49	9.72	9.98	10.23	10.53	10.19	10.07	10.48
	Sample pixel 2		12.42	13.15	12.08	11.70	11.88	12.01	12.38	11.22	11.94	12.16

640
641
642
643



644
 645 **Figure A1:** OMI SO₂ VCD for the four volcanic cases: (a) Kasatochi on August 10th, 2008, (b)
 646 Kelud on February 14th, 2014, (c) Calbuco on April 24th, 2015 and (d) Raikoke on June 24th,
 647 2019. In these maps, only pixels with SO₂ > 10 DU are shown.

648
 649
 650 **Data availability.** OMI SO₂ L1 and L2 data can be accessed via the Goddard Earth Sciences Data and
 651 Information Services Center (GES DISC) at <https://earthdata.nasa.gov/eosdis/daacs/gesdisc>. IASI SO₂
 652 LH data is available via the IASI AERIS portal <https://iasi.aeris-data.fr/>. NASA CALIPSO data can be
 653 downloaded from <https://www-calipso.larc.nasa.gov/> and images can be found at https://www-calipso.larc.nasa.gov/products/lidar/browse_images/production. TROPOMI L2 SO₂ data can be obtained
 654 at <https://s5phub.copernicus.eu/dhus/#/home> while the LH is experimental and is not yet publicly
 655 available online. The results of OMI SO₂ layer height retrieval presented in this study can be obtained
 656 from the author by request.

657
 658
 659 **Author contributions.** NF wrote the manuscript and performed most computational and model work in this
 660 study. The project was conceived and overseen by CL and NK. DL and PH provided the TROPOMI SO₂ LH
 661 retrieval data and input on the comparisons in the paper. PH also offered support relating to the machine
 662 learning aspect of the study. RS is the original developer of the LIDORT-RRS code and provided related
 663 support, as well as input to the relevant sections of the manuscript. RD is an advisor of NF and provided
 664 additional input to the paper and was involved in project planning.

665

666 **Competing interests.** The authors declare that they have no conflict of interest.

667

668 **Acknowledgements.** We would like to acknowledge the NASA Earth Science Division (ESD) Aura Science
669 Team program for funding of the OMI SO₂ product development and analysis (Grant # 80NSSC17K0240).

670 OMI is a Dutch/Finish contribution to the NASA Aura mission. The OMI project is managed by the Royal
671 Meteorological Institute of the Netherlands (KNMI) and the Netherlands Space Agency (NSO).

672

673

674 **References.**

675

676 Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O. C., Vogel, A., Hartmann, M.,
677 Bovensmann, H., Frerick, J., and Burrows, J. P.: Measurements of molecular absorption spectra with the
678 SCIAMACHY pre-flight model: Instrument characterization and reference data for atmospheric remote
679 sensing in the 230-2380 nm region, *J. Photochem. Photobiol. A: Chem.* 157, 167-184,
680 doi: 10.1016/S1010-6030(03)00062-5, 2003.

681

682 Carn, S. A., A. J. Krueger, N. A. Krotkov, K. Yang, and K. Evans. “Tracking Volcanic Sulfur Dioxide
683 Clouds for Aviation Hazard Mitigation, *Natural Hazards* 51 (2): 325–343, doi:10.1007/s11069-008-9228-
684 4, 2009.

685

686 Carn, S. A., Fioletov, V. E., McLinden, C. A., Li, C., Krotkov, N. A., *Scientific Reports*, 7, 44095,
687 doi: 10.1038/srep44095, 2017.

688

689 Chance, K., and R. L. Kurucz. An improved high-resolution solar reference spectrum for Earth’s
690 atmosphere measurements in the ultraviolet, visible, and near infrared, *J. Quant. Spectrosc. Radiat.*
691 *Transfer*, 111, 1289–1295, doi:10.1016/j.jqsrt.2010.01.036, 2010.

692

693 Clarisse, L., P. F. Coheur, A. J. Prata, D. Hurtmans, A. Razavi, T. Phulpin, J. Hadji-Lazaro, and C.
694 Clerbaux. “Tracking and Quantifying Volcanic SO₂ with IASI, the September 2007
695 Eruption at Jebel at Tair.” *Atmospheric Chemistry & Physics* 8: 7723–7734. doi:10.5194/acp-8-
696 7723-2008, 2008.

697

698 Clarisse, L., Coheur, P. F., Theys, N., Hurtmans, D., and Clerbaux, C.: The 2011 Nabro eruption, a SO₂
699 plume height analysis using IASI measurements, *Atmos. Chem. Phys.*, 14, 3095–3111,
700 <https://doi.org/10.5194/acp-14-3095-2014>, 2014.

701

702 Daumont, D., Brion, J., Charbonnier, J., and Malicet, J.: Ozone UV spectroscopy. I: Absorption cross-
703 sections at room temperature, *J. Atmos. Chem.*, 15, 145 – 155, doi:10.1007/BF00053756, 1992.

704

705 Efremenko, D. S., Loyola R., D. G., Hedelt, P., and Spurr, R. J. D.: Volcanic SO₂ plume height retrieval
706 from UV sensors using a full-physics inverse learning machine algorithm,
707 *International Journal of Remote Sensing*, 38, 1–27, <https://doi.org/10.1080/01431161.2017.1348644>,
708 2017.

709

710 Fioletov, V. E., McLinden, C. A., Krotkov, N., and Li, C.: Lifetimes and emissions of SO₂ from point
711 sources estimated from OMI, *Geophys. Res. Lett.*, 42, 1969-
712 1976, <https://doi.org/10.1002/2015GL063148>, 2015.
713

714 Guffanti, M., T. J. Casadevall, and K. Budding. Encounters of aircraft with volcanic ash clouds: A
715 compilation of known incidents, 1953-2009, Tech. rep., U. S. Geological Survey, Data Series 545, ver.
716 1.0. [Available at <http://pubs.usgs.gov/ds/545/>, 2010.
717

718 Halton, J. H.: On the Efficiency of Certain Quasi-Random Sequences of Points in Evaluating Multi-
719 Dimensional Integrals. *Numerical Mathematical* 2 (1), 84–90, doi:10.1007/BF01386213, 1960.
720

721 Hedelt, P., Efremenko, D. S., Loyola, D. G., Spurr, R., and Clarisse, L.: SO₂ Layer Height retrieval from
722 Sentinel-5 Precursor/TROPOMI using FP_ILM, *Atmos. Meas. Tech.*, 12, 5503–5517, 2019
723 <https://doi.org/10.5194/amt-12-5503-2019>, 2019
724

725 Lee, C., R. V. Martin, A. Van Donkelaar, R. R. Hanlim Lee, J. C. H. Dickerson, N. Krotkov, A. Richter,
726 K. Vinnikov, and J. J. Schwab.: SO₂ Emissions and Lifetimes: Estimates from Inverse Modeling Using in
727 Situ and Global, Space-Based (SCIAMACHY and OMI) Observations, *Journal of Geophysical Research:*
728 *Atmospheres* 116: (D6): n/a–n/a. D06304. doi:10.1029/2010JD014758, 2011.
729

730 Levelt, P. F., Van Den Oord, G. H. J., Dobber, M. R., Mälkki, A., Visser, H., De Vries, J., Stammes, P.,
731 Lundell, J. O. V., and Saari, H.: The Ozone Monitoring Instrument, *IEEE Trans. Geosci. Remote Sens.*,
732 44, 1093–1101, 2006b.
733

734 Li, C., Joiner, J., Krotkov, N. A., and Bhartia, P. K.: A fast and sensitive new satellite SO₂ retrieval
735 algorithm based on principal component analysis: Application to the ozone monitoring instrument,
736 *Geophys. Res. Lett.*, 40, 6314–6318, doi:10.1002/2013GL058134, 2013.
737

738 Li, C., Krotkov, N. A., Carn, S., Zhang, Y., Spurr, R. J. D., and Joiner, J.: New-generation NASA Aura
739 Ozone Monitoring Instrument (OMI) volcanic SO₂ dataset: algorithm description, initial results, and
740 continuation with the Suomi-NPP Ozone Mapping and Profiler Suite (OMPS), *Atmos. Meas. Tech.*, 10,
741 445– 458, <https://doi.org/10.5194/amt-10-445-2017>, 2017.
742

743 Kristiansen, N. I., Prata, A. J., Stohl, A., and Carn, S. A.: Stratospheric volcanic ash emissions from the
744 13 February 2014 Kelut eruption, *Geophys. Res. Lett.*, 42, 588–596, doi:10.1002/2014GL062307, 2015.
745

746 Lee, C., R. V. Martin, A. Van Donkelaar, R. R. Hanlim Lee, J. C. H. Dickerson, N. Krotkov, A. Richter,
747 K. Vinnikov, and J. J. Schwab.: SO₂ Emissions and Lifetimes: Estimates from Inverse Modeling Using in
748 Situ and Global, Space-Based (SCIAMACHY and OMI) Observations, *Journal of Geophysical Research:*
749 *Atmospheres* 116: (D6): n/a–n/a. D06304. doi:10.1029/2010JD014758, 2011.
750

751 Loyola, D. G., M. Pedernana, and S. Gimeno Garcia.: Smart Sampling and Incremental
752 Function Learning for Very Large High Dimensional Data. *Neural Networks* 78: 75–87.
753 doi:10.1016/j.neunet.2015.09.001, 2016.
754

755 Loyola, D. G., Xu, J., Heue, K.-P., and Zimmer, W.: Applying FP_ILM to the retrieval of geometry-
756 dependent effective Lambertian equivalent reflectivity (GE_LER) daily maps from UVN satellite
757 measurements, *Atmos. Meas. Tech.*, 13, 985–999, <https://doi.org/10.5194/amt-13-985-2020>, 2020.

758
759 Ludewig, A., Kleipool, Q., Bartstra, R., Landzaat, R., Leloux, J., Loots, E., Meijering, P., van der Plas, E.,
760 Rozemeijer, N., Vonk, F., and Veeffkind, P.: In-flight calibration results of the TROPOMI payload on-
761 board the Sentinel-5 Precursor satellite, *Atmos. Meas. Tech.*, 13, 3561–3580, [https://doi.org/10.5194/amt-](https://doi.org/10.5194/amt-13-3561-2020)
762 13-3561-2020, 2020.
763
764 Kristiansen, N. I., Prata, A. J., Stohl, A., and Carn, S. A.: Stratospheric volcanic ash emissions from the 13
765 February 2014 Kelut eruption, *Geophys. Res. Lett.*, 42, 588–596, doi:10.1002/2014GL062307, 2015.
766
767 McCormick, M. P., L. W. Thomason, and C. R. Trepte.: Atmospheric Effects of the Mt Pinatubo
768 Eruption, *Nature* 373: 399–404. doi:10.1038/373399a0., 2015.
769
770 Nowlan, C. R., X. Liu, K. Chance, Z. Cai, T. P. Kurosu, C. Lee, and R. V. Martin.: Retrievals of Sulfur
771 Dioxide from the Global Ozone Monitoring Experiment 2 (GOME-2) Using an Optimal Estimation
772 Approach: Algorithm and Initial Validation, *Journal of Geophysical Research: Atmospheres* 116 (D18):
773 n/a–n/a. D18301. doi:10.1029/2011JD015808, 2011.
774
775 Rix, M., P. Valks, N. Hao, D. Loyola, H. Schlager, H. Huntrieser, A. Flemming, U. Koehler, U.
776 Schumann, and A. Inness.: Volcanic SO₂, BrO and Plume Height Estimations Using GOME-2 Satellite
777 Measurements during the Eruption of Eyjafjallajökull in May 2010, *Journal of Geophysical Research*
778 (Atmospheres) 117: D00U19. doi:10.1029/2011JD016718, 2012.
779
780 Schenkeveld, V. M. E., Jaross, G., Marchenko, S., Haffner, D., Kleipool, Q. L., Rozemeijer, N. C.,
781 Veeffkind, J. P., and Levelt, P. F.: In-flight performance of the Ozone Monitoring Instrument, *Atmos.*
782 *Meas. Tech.*, 10, 1957–1986, <https://doi.org/10.5194/amt-10-1957-2017>, 2017.
783
784 Schmidt, A.; Witham, C.S.; Theys, N.; Richards, N.A.D.; Thordarson, T.; Szpek, K.; Feng, W.; Hort,
785 M.C.; Woolley, A.M.; Jones, A.R.; Redington, A.L.; Johnson, B.T.; Hayward, C.L.; Carslaw, K.S.:
786 Assessing hazards to aviation from sulfur dioxide emitted by explosive Icelandic eruptions. , *Journal of*
787 *Geophysical Research D: Atmospheres*, Vol. 119, Issue 24, 14180-14196, doi: 10.1002/2014JD022070,
788 2014.
789
790 Spurr, R., de Haan, J., van Oss, R., and Vasilkov, A.: Discreteordinate radiative transfer in a stratified
791 medium with first-order rotational Raman scattering, *J. Quant. Spectrosc. Ra.*, 109, 404–425,
792 <https://doi.org/10.1016/j.jqsrt.2007.08.011>, 2008.
793
794 Torres, O., Bhartia, P. K., Jethva, H., and Ahn, C.: Impact of the ozone monitoring instrument row
795 anomaly on the long-term record of aerosol products, *Atmos. Meas. Tech.*, 11, 2701–2715,
796 <https://doi.org/10.5194/amt-11-2701-2018>, 2018.
797
798 Vernier, J.-P., Fairlie, T. D., Deshler, T., Natarajan, M., Knepp, T., Foster, K., Wienhold, F. G., Bedka, K.
799 M., Thomason, L., and Trepte, C.: In situ and space-based observations of the Kelud volcanic plume: The
800 persistence of ash in the lower stratosphere, *J. Geophys. Res. Atmos.*, 121, 11104–11118,
801 <https://doi.org/10.1002/2016JD025344>, 2016.
802
803 von Glasow, R., Bobrowski, N., and Kern, C.: The effects of volcanic eruptions on atmospheric
804 chemistry, *Chem. Geol.*, 263, 131–142, <https://doi.org/10.1016/j.chemgeo.2008.08.020>, 2009.
805

806 Xu, J., Schüssler, O., Loyola R., D., Romahn, F., and Doicu, A.: A novel ozone profile shape retrieval
807 using Full-Physics Inverse Learning Machine (FP_ILM)., IEEE J. Sel. Topics Appl. Earth Observ.
808 Remote Sens., 10, 5442–5457, <https://doi.org/10.1109/JSTARS.2017.2740168>, 2017.
809

810 Yang, K., N. A. Krotkov, A. J. Krueger, S. A. Carn, P. K. Bhartia, and P. F. Levelt.: Retrieval of large
811 volcanic SO₂ columns from the Aura Ozone Monitoring Instrument: Comparison and limitations, J.
812 Geophys. Res., 112, D24S43, doi:10.1029/2007JD008825, 2007.
813

814 Yang, K., X. Liu, N. A. Krotkov, A. J. Krueger, and S. A. Carn.: Estimating the altitude of volcanic sulfur
815 dioxide plumes from space borne hyper-spectral UV measurements, Geophys. Res. Lett., 36, L10803,
816 doi:10.1029/2009GL038025, 2009.
817

818 Yang, K., P. K. Xiong Liu, N. A. Bhartia, S. A. Krotkov, E. J. Carn, A. J. Hughes, R. J. Krueger, D.
819 Spurr, and S. G. Trahan.: Direct Retrieval of Sulfur Dioxide Amount and Altitude from
820 Spaceborne Hyperspectral UV Measurements: Theory and Application, Journal of Geophysical
821 Research: Atmospheres 115: D2. doi:10.1029/2010JD013982, 2010.
822

823 Young, A. T.: Rayleigh Scattering. Applications Optical 20 (4): 533–535. doi:10.1364/
824 AO.20.000533, 1981.
825
826