A multi-sensor satellite-based archive of the largest SO₂ volcanic eruptions since 2006

Pierre-Yves Tournigand[1], Valeria Cigala[2], Elzbieta Lasota[1], Mohammed Hammouti[3], Lieven Clarisse[4], Hugues Brenot[5], Fred Prata[6], Gottfried Kirchengast[7], Andrea K. Steiner[7], Riccardo Biondi[1]

5   [1]Dipartimento di Geoscienze, Università degli Studi di Padova, Italy
[2]Ludwig-Maximilians-Universität München, Germany
[3]Politecnico di Milano, Italy
[4]Spectroscopy, Quantum Chemistry and Atmospheric Remote Sensing (SQUARES), Université Libre de Bruxelles, Belgium
10   [5]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
[6]AIRES Pty Ltd: Mt Eliza, Victoria, Australia
[7]Wegener Center for Climate and Global Change, University of Graz, Austria

Correspondence to: Riccardo Biondi (riccardo@biondiriccardo.it)

Abstract. We present a multi-sensor archive collecting spatial and temporal information about volcanic SO₂ clouds
15   generated by the eleven largest eruptions of this century. The detection and monitoring of volcanic clouds is an important topic for aviation management, climate issues and weather forecast. Several papers have been published focusing on single events, but not any archive is available at the moment to be used as background for future studies. We archived and collocated the SO₂ vertical column density estimations from three different instruments (AIRS, IASI and GOME-2), the atmospheric parameters vertical profiles from the Global Navigation Satellite Systems (GNSS)
20   Radio Occultations (RO) and the vertical backscatter from the Cloud-Aerosol LIdar with Orthogonal Polarization (CALIOP). We additionally provide information about the cloud top height from three different algorithms and the atmospheric anomaly due to the presence of the cloud. The dataset consists of 223 days monitored with SO₂ clouds, collocated with 56675 backscatter profiles and 70126 radio occultation profiles. The modular structure of the archive allows an easy collocation of the different datasets according to the users' needs and the cross-comparison of the
25   datasets shows the high consistency of the parameters estimated with different sensors and algorithms. The data described here will be published with a DOI after final acceptance of this manuscript (Tournigand et al., 2020, http://doi.org/10.5880/fidgeo.2020.016). During the discussion period, the data are accessible via this temporary link: http://pmd.gfz-
potsdam.de/panmetaworks/review/0f85d699707efcdc567765bd0dafaaadf94b6df5a531f310167f7e974ea803bf.

30   ## 1. Introduction

Volcanoes around the world are a constant source of gaseous emissions. Both passively, during quiescent times, and actively, during eruptions (Robock, 2015; Carn et al., 2017). The most abundant gas species emitted are water (H₂O), carbon dioxide (CO₂) and sulfur dioxide (SO₂), while less abundant ones include hydrogen (H₂), hydrogen sulfur (H₂S), hydrochloric acid (HCl) and carbon monoxide (CO) (Williams-Jones and Rymer, 2015). Once emitted into the
35   atmosphere, some of these gases can react and transform into aerosols, for example, SO₂ transforms into H₂SO₄. As with all volcanic eruption products, the gases emitted and the related aerosols can pose hazards to people as well as the environment (Williams-Jones and Rymer, 2015). Moreover, they can be responsible for regional and global climatic effects, depending on the latitude of the volcano, the altitude reached by the eruptive column and consequent cloud (Robock, 2000; Robock, 2015; Williams-Jones and Rymer, 2015; Carn et al., 2016). In terms of global climatic
40   impact, SO₂ injections in the stratosphere are of the greatest significance. The reason is that in the stratosphere the gas and subsequent aerosol can remain suspended for months to years and hence transported around the globe affecting the absorption of both short- and longwave radiation (Robock, 2015). The duration and spatial spreading of emitted gases and aerosols in the atmosphere also depend on the erupted mass of volcanic material and the duration of the emission.

According to the Global Volcanism Program (GVP) of the Smithsonian Institute, an average of 55 to 88 eruptions (excluding permanent and semi-permanent degassing) has occurred per year since 1994 worldwide. The eruptions display variability in their eruptive style (e.g., effusive or explosive), magma composition, the energy of the eruption, amount, type and size of the ejected material. To compare and characterize different eruptive events in an objective manner, the Volcanic Explosivity Index (VEI) was created. The VEI was introduced in 1982, drawing inspiration from the Richter's scale for earthquakes' magnitude, to provide a relative, semi-quantitative measure of the explosiveness of volcanic eruptions by Newhall and Self (1982). The VEI classification, divided into categories from 0 to 8, is based mainly on measures of magnitude, in terms of total ejecta volume, and/or intensity, in terms of mass flux or eruption plume height, depending on data availability (Newhall and Self, 1982; Houghton et al., 2013). The VEI index has its limitations, nevertheless, it is being used extensively to provide an eruption descriptor that is understandable by researchers and policy-makers alike (Houghton et al., 2013).

The size of an eruption can, however, be significantly different when considering gas emissions only (Carn et al., 2016). Thus, a different eruption size classification can be outlined by the sulfur input into the atmosphere (Schnetzler et al., 1997; Robock, 2015; Carn et al., 2016). In 1997, Schnetzler et al. proposed the Volcanic Sulfur Dioxide Index (VSI) based on SO_2 retrievals performed with the Total Ozone Mapping Spectrometer (TOMS) onboard the Nimbus 7 satellite. Medium VEI eruptions, e.g. VEI 4 events, can be characterized by the emission of a large quantity of tephra with respect to the quantity of SO_2 (e.g. the 2008 Chaiten eruption) or by the emission of a larger amount of SO_2 than tephra (e.g. the 2011 Nabro eruption) as shown in Carn et al. (2016). Using an improved SO_2 emissions retrieval approach and including less energetic events, Carn et al. (2016) found a broader range of SO_2 emission per VEI and a weaker first-order correlation. These findings suggest that the intensity and volcanic plume altitude are more relevant parameters for consideration in modelling SO_2 emissions and their climate impact (Robock, 2000; Carn et al., 2016).

The above-mentioned GVP provides the most extensive catalogue of historical eruptions with information related to both volcanoes and their eruptions. This catalogue is a firsthand source of information when starting an investigation of a given volcano or a given style of eruption. Similarly, the Large Magnitude Explosive Volcanic Eruptions (LaMEVE) dataset includes data such as the magnitude of the event, the bulk volume, the tephra fallout volume, column height of Quaternary (from 2.58 Ma ago to present) eruptions with VEI ≥ 4 (Brown et al., 2014). Other datasets are available including data on geochemical composition (e.g. Turner et al., 2009), or acoustic acquisitions (e.g. Fee et al., 2014), even though these types of database are generally limited to a specific volcano or a specific time window (e.g. de Moor et al., 2017).

Previous papers focusing on single or a few eruptions are based on personal data collections or project collaborations and this makes it difficult for data comparison and studies with new techniques or algorithms. The eruptions of Okmok and Kasatochi in 2008 were the focus of a special issue (JGR Atmospheres, 2018) collecting 27 papers studying the cloud with all the available remote sensing platforms and algorithms. Sarychev Peak 2009 volcanic cloud was also well studied (e.g., Carn and Lopez, 2011; Kravitz et al., 2011; Rybin et al., 2011; Doeringer et al., 2012). The Eyjafjallajokull 2010 eruption affected the economy and social life of Europe and beyond, changing the rules of air traffic management and the volcanic cloud was subject of a number of studies (e.g., Flentje et al., 2010; Marenco et al., 2011; Prata and Prata, 2012; Stohl et al., 2012). Grimsvotn 2011 was a quite interesting eruption from a scientific point of view because the SO_2 cloud was separated from the ash cloud (Moxnes et al., 2014) and different researchers studied it (e.g., Flemming and Inness, 2013; Marzano et al., 2013; Cook et al., 2014; Prata et al., 2017). The Nabro 2011 eruption was the subject of an interesting discussion regarding the direct intrusion to the stratosphere of the volcanic cloud (e.g., Bourassa et al., 2012; Clarisse et al., 2014; Fromm et al., 2014; Biondi et al., 2017) and the Puyehue Cordon Caulle, erupting in the same period, was of interest because the cloud moved around the globe (e.g., Bignami et al., 2014; Griessbach et al., 2014; Theys et al., 2014; Biondi et al., 2017). However, other volcanic clouds, such as the ones produced by Merapi 2010 (Picquout et al., 2013), Tolbachik 2012 (Telling et al., 2015), Kelut 2014 (Kristiansen et al., 2015; Vernier et al., 2016) and Calbuco 2015 (Marzano et al., 2018; Lopes et al., 2019) were not studied in depth.

Considering SO_2 emissions, several datasets and inventories are available and updated in the course of the years. In recent times, Ge et al. (2016) compiled an inventory for SO_2 emissions in the time frame 2005-2012 retrieved by the Ozone Monitoring Instrument (OMI) onboard the Aura satellite. Carn et al. (2017) implemented it including retrievals

made until 2015 and emissions related to passive degassing in addition to the ones from main eruptive events. The most updated and complete dataset up to now is the Multi-Satellite Volcanic Sulfur Dioxide Database Long-Term L4 Global (MSVOLSO2L4) compiled by Carn (2019). The dataset provides "SO₂ mass loadings for all significant global volcanic eruptions detected from space since October 1978" to present (Carn, 2019). The MSVOLSO2L4 includes ancillary information about the volcanoes, such as the name and location of the volcano, as well as information about the eruptions, for example, start and end date, and VEI. This information is retrieved from the GVP database. The dataset also reports the observed plume altitude (in kilometres) where known. Otherwise, an estimated plume altitude above vent depending on eruption type, and the measured SO_2 mass in kilotons (= 1000 metric tons) is provided (Carn et al., 2016; Carn, 2019).

In this study, we present the first database predicated on satellite data, reporting:

1. SO_2 retrievals from Atmospheric Infrared Sounder (AIRS), Infrared Atmospheric Sounding Interferometer (IASI), and the Global Ozone Monitoring Experiment 2 (GOME-2). The data from the three sensors provide horizontal and temporal information on SO_2 concentrations;
2. SO_2 cloud altitude estimations from IASI, the Cloud-Aerosol LIdar with Orthogonal Polarization (CALIOP) backscatter and the Global Navigation Satellite System (GNSS) Radio Occultation (RO);
3. the cloud aerosol subtype information from CALIOP;
4. atmospheric properties such as temperature, pressure, humidity, from GNSS RO profiles.

The information is provided for eruptions classified by GVP as VEI 4 or larger, with an SO_2 total mass loading larger than 0.05 Tg and that occurred from 2006 to 2016 since not any eruption in the periods 2016-2019 was yet classified as VEI 4 at the time of the archive preparation. The selected volcanoes and relative eruptions are (Table 1): Okmok 2008; Kasatochi 2008; Sarychev 2009; Eyjafjallajokull 2010; Merapi 2010; Grimsvotn 2011; Nabro 2011; Puyehue-Cordon Caulle 2011; Tolbachik 2012; Kelut 2014; Calbuco 2015. These are the most significant eruptions over the period 2006-2018.

To the best of our knowledge, there is no current unique database collecting SO_2 maps and volcanic cloud altitude estimations from several instruments, cloud aerosols subtypes and atmospheric properties for explosive eruptions. Accurate knowledge on volcanic SO_2 clouds concentration and altitude as well as, their spatial and temporal evolution, is of crucial importance in the investigation of an eruption climatic impact. Thus, we believe that the database presented here will help current and future investigations as well as support the development of more accurate retrievals methodologies.

2. Instrument description

A summary of the instruments' characteristics together with parameters provided in this work and references to the algorithms are reported in Table 2. In this archive, each sensor measuring SO_2 amounts measures the partial column density, due to their own limitations (Brenot et al., 2014). We here use the terms vertical column density (VCD) to refer to this partial column density.

2.1. AIRS

The Atmospheric Infrared Sounder (AIRS) is a cross-track hyperspectral instrument onboard the polar-orbiting satellite Aqua launched in June 2002 with ascending equator crossing local time at 13:30. AIRS completely covers the full globe two times per day with a swath of 1650 km and spatial resolution of 13.5 x 13.5 km² at nadir and 41 x 21 km² at high latitudes (Susskind et al., 2003). The SO_2 pixels are identified using infrared channels centered at the 7.3 µm absorption peak relying on the correlation between the measured spectrum and a reference spectral shape. The amount of SO_2 in each pixel is computed by a least-squares procedure based on an off-line radiative transfer model. This technique performs well for SO_2 reaching high tropospheric altitudes or the stratosphere where the water vapor content is negligible. Comparisons with other techniques (Carn et al., 2016; Carn et al., 2017) show an agreement within 10–30% and a typical retrieval error for a single AIRS pixel of about 6 Dobson Unit (DU) (Prata and Bernardo, 2007).

2.2.   IASI

140   The Infrared Atmospheric Sounding Interferometer (IASI) is a Fourier transform instrument onboard the near-polar sun-synchronous orbiting satellites Metop-A and Metop-B, respectively, launched in October 2006 and September 2012 with ascending equator crossing local time at 9:30. IASI completely covers the full globe two times per day with a swath of 2200 km and a spatial resolution of 12 km at nadir (Clerbaux et al., 2009). The SO_2 retrieval is based on a brightness temperature difference in the SO2 v_3 band (Clarisse et al., 2012) which is converted to SO_2 concentration
145   integrated along the vertical axis the Vertical Column Density (VCD) using look-up tables and operational profiles of pressure, temperature and humidity. The retrieval of VCD assumes that all SO_2 is located at particular atmospheric layers (5, 7, 10, 13, 16, 19, 25 or 30 km above sea level) providing different estimations at different altitudes. Then a second algorithm (Clarisse et al., 2014) is applied to compute the SO_2 cloud altitude with an accuracy of about 2 km for plumes below 20 km.

150   ### 2.3.   GOME

The Global Ozone Monitoring Experiment 2 (GOME-2) is an ultraviolet-visible spectrometer, on board of the Metop-A and Metop-B satellites, measuring solar light backscattered by the atmosphere or reflected by the Earth in nadir-viewing geometry with a swath of 1920 km and spatial resolution of 40x80 km at nadir (Munro et al., 2006). The SO_2 VCD retrieval (Rix et al., 2012) is based on the strong SO_2 absorption between 240 and 400 nm and uses a differential
155   optical absorption spectroscopy technique (Platt and Stutz, 2008). All measurements in the wavelength ranging from 315 to 326 nm are fitted to laboratory absorption data of SO_2 and converted to VCD with an air mass factor from radiative transfer models assuming hypothetical atmospheric layers representative of different scenarios of emissions. The SO_2 VCD is provided for 3 atmospheric layers representative of different scenarios of emissions: low troposphere (~2.5 km above the surface), upper troposphere (∼6 km) and lower stratosphere (∼15 km).

160   ### 2.4.   CALIOP

The Cloud-Aerosol LIdar with Orthogonal Polarization (CALIOP) is an instrument onboard the polar-orbiting Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO). To estimate the volcanic cloud (VC) top altitude and validate the cloud top estimation from GNSS RO, we have used the Level 1 total attenuated backscatter at 532 nm (CAL_LID_L1 Version 4) with a vertical resolution of 60 m and horizontal resolution of 1 km between 10
165   and 20 km of altitude, and a vertical resolution of 180 m and horizontal resolution of 1,67 km above 20 km (Winker et al., 2009). To extract the corresponding aerosol type we have used the Level 2 Vertical Feature Mask (Winker et al., 2009) product version 4.20.

2.5.   GNSS RO

The Global Navigation Satellite Systems (GNSS) Radio Occultation (RO) is an active limb sounding technique which
170   uses the signal transmitted by a GNSS satellite and received by a Low Earth Orbit satellite. The signal, travelling through the atmospheric layers, is bent according to the different refractivity of each layer and thus provides information about the vertical structure of the troposphere and stratosphere (Kursinski et al., 1997). The vertical resolution of the RO ranges between 100 m in the upper troposphere to about 500 m in the lower stratosphere at low/mid-latitudes (Zeng et al., 2019), while the horizontal resolution can range from about 50 km in the troposphere
175   to 200-300 km in the stratosphere (Kursinski et al., 1997). We use for this archive the RO bending angle, refractivity, temperature, pressure and specific humidity profiles processed by the Wegener Center for Climate and Global Change (WEGC) with the Occultation Processing System (OPS) version 5.6 (EOPAC Team, 2019). We also provide the bending angle anomaly which is proven to be an efficient parameter to understand the impact of the VC on the atmospheric structure (Biondi et al., 2017; Cigala et al., 2019; Stocker et al., 2019). The way this anomaly is computed
180   is detailed in section 4.1. The RO profiles are obtained using a combination of geometric optics and wave optics retrieval (Angerer et al., 2017), with transition below the tropopause. The retrieval is based on orbit information and amplitude and phase data from the University Corporation of Atmospheric Research/COSMIC Data Analysis and

185 Archive Center (UCAR/CDAAC) collected from the following RO missions: the CHAllenging Minisatellite Payload (CHAMP) (Wickert et al., 2001), the Satélite de Aplicaciones Científicas (SAC-C) (Hajj et al., 2004), the Gravity Recovery And Climate Experiment (GRACE-A) (Beyerle et al., 2005), the FORMOSAT-3/COSMIC (Anthes et al., 2008), and the EUMETSAT/METOP missions (Luntama et al., 2008). The accuracy of the GNSS RO is 0.2 K in terms of temperature and 0.1% in terms of refractivity and the data from the different mission are very consistent (Scherllin-Pirscher et al., 2011), so there is no need of inter-calibration or homogenization (Foelsche et al., 2011).

3. Data description

190 The archive (Tournigand et al., 2020) consists of two sets of files, the daily files and the eruption files, i.e., one file per eruption including all collected information. Thus, for each eruption listed in Table 1 the user can choose to access one single day or to the whole eruptive period depending on the user's demand. The amount of days covered by the archive for each eruption depends on the SO_2 detection availability from AIRS, IASI and GOME-2. Each file is in NetCDF-4 format and file names are self-explanatory with daily files following the format 195 *volcanoname_year_month_day* and the eruption files following the format *volcanoname*. As an example, a user who wishes to access all available data corresponding to the Sarychev volcano on 12 of June 2009 will have to look for the file *sarychev_2009_06_12.nc*. The organization of both file types is described hereafter for each instrument.

3.1. AIRS

AIRS data are organized in the same way in both file types. It consists of 4 variables namely AIRS_lat, AIRS_lon, 200 AIRS_date and AIRS_SO2 respectively containing the latitude (°N), longitude (°E), date and time (POSIX time, number of seconds elapsed since 00:00:00 UTC 1st of January 1970) of each granule and their SO_2 VCD (DU). The variables AIRS_lat, AIRS_lon and AIRS_SO2 are matrices with columns corresponding to the different granules. By selecting one column, the user can find each data point of the corresponding granule. The AIRS_date variable, on the other hand, is a line vector with elements reporting the date and time of each granule. Only data points with SO_2 values 205 higher than 0 DU have been included in the archive thus explaining the different amount of points from one granule to another.

3.2. IASI

IASI data are organized in the same way for both file types and are composed of 5 variables, IASI_lat, IASI_lon, IASI_date, IASI_SO2 and IASI_height respectively containing the latitude (°N), longitude (°E), date and time (POSIX 210 time), SO_2 (DU) and cloud altitude (m). The date variable consists of a line vector with elements corresponding to each granule. Similarly, the other variables are matrices with columns corresponding to the different granules and lines to the data points of the given granule having an SO_2 content higher than 0 DU.

3.3. GOME-2

GOME-2 data's organization is identical in both file types. GOME-2 is composed of 6 variables, GOME_lat, 215 GOME_lon, GOME_date, GOME_SO2_1, GOME_SO2_2, GOME_SO2_3 respectively corresponding to the latitude (°N), longitude (°E), date and time (POSIX time), low troposphere SO_2 vertical column density (DU), mid-troposphere SO_2 vertical column density (DU) and low stratosphere SO_2 vertical column density (DU). As for AIRS and IASI, the date variable corresponds to a line vector providing each granule's date. The rest of the variables correspond to matrices with granules also separated in columns and the data points of those granules distributed in lines. In the case 220 of GOME-2, only data points having their three SO_2 vertical columns contents higher than 0 DU were included. Pixel with high Solar Zenith Angle (SZA) has also been filtered.

3.4.   CALIOP

CALIOP's section of the archive contains 6 variables, CALIOP_lat, CALIOP_lon, CALIOP_date, CALIOP_filename, CALIOP_height, CALIOP_type respectively corresponding to the latitude (°N), longitude (°E),
225   date and time (POSIX time), name of CALIOP file, estimated VC top altitude (m) and aerosol type. The latitude, longitude, date and file name variables are column vectors with each line corresponding to latitude, longitude and date data of the designated file in CALIOP_filename variable. CALIOP_height variable contains all the cloud top altitude estimations based on CALIOP L1 532 nm version 4.10 backscatter product. This variable is a matrix with each line corresponding to a CALIOP file and the three columns indicating to which instrument the CALIOP file is collocated
230   with, AIRS (column 1), IASI (column 2) or GOME (column 3).

CALIOP_type variable contains the type of aerosol retrieved from the L2 Vertical Feature Mask (VFM) version 4.20. The L2 VFM CALIOP product subdivides the aerosols into 10 types. Four of those types are of interest for this archive type 2, 6, 9 and 10 respectively corresponding to dust, elevated smoke, volcanic ash and sulfate/other. This variable is subdivided into three columns corresponding to three levels of altitude -0.5 to 8.2 km, 8.2 to 20.2km and 20.2 to
235   30.1km. In each altitude level, the presence of one or several cloud types is indicated by the presence of their reference number.

3.5.   GNSS RO

GNSS RO data are organized in different ways in the two file types. In the daily files, the RO data are separated in different variables according to the instrument they are collocated with (AIRS, IASI or GOME-2). For each set of RO
240   data collocated with a given instrument 10 variables are provided, latitude (°N), longitude (°E), date (POSIX time), bending angle (rad), bending angle anomaly (%), temperature (K), pressure (Pa), refractivity (N-unit), specific humidity (kg.kg^{-1}) and volcanic cloud top altitude (m). Each variable is a matrix with each column corresponding to a RO profile and lines to the latitude dimension. Only the variable containing volcanic cloud altitude is a line vector with each element corresponding to a different RO profiles.

245   In the files containing the whole eruptive period, the RO data are not separated according to the instrument they are collocated with but compiled all together. Thus, the same 10 variables are provided as in daily files, each containing the totality of the RO profiles.

4.   Quality control and data processing

4.1.   RO data

250   The RO profiles included in this archive are collocated spatially at ±0.2° and temporally at ±12h with data points from the volcanic aerosol maps provided by AIRS, IASI, and GOME-2 acquisitions.

4.1.1 Climatology

The RO reference climatology for each area of interest is calculated based on 5° latitude bands using our dataset of RO profiles covering a period from 2001 to 2017. The averaging of all available RO profiles present within each
255   latitude band provides the RO reference climatology.

4.1.2 Anomaly calculation

The bending angle (BA) anomaly integrated into this archive is calculated by subtracting the RO reference climatology profile from the individual RO BA profile and normalized with respect to the reference climatology profile (Eq 1), following the methodology described in Biondi et al. (2017). The resulting anomaly displays variations when the
260   bending angle differs from the climatology. Such variations indicate a change of atmospheric properties and are used

to identify related atmospheric features. The presence of volcanic clouds in the atmosphere generates a prominent peak in the BA anomaly profile.

$$\alpha = \left(\frac{(BA - BA_{clim})}{BA_{clim}}\right) \cdot 100 \qquad (1)$$

With **α** the bending angle anomaly, BA the individual bending angle profile and BA_{clim} the BA climatology profile.

4.1.3 Peak detection

The peak detection of bending angle anomaly profiles was automatically done using a customized Matlab algorithm. This algorithm, further developed after Cigala et al. (2019), identifies all the peaks displaying a variation larger than 4.5% in the bending angle anomaly profile with respect to local minimums. Only the peaks having their maximum value between 10 and 22 km of altitude are kept. Peaks vertically spreading over more than 8 km have been removed. Finally, amongst the remaining peaks, the lowest altitude one is selected as a cloud top altitude.

4.2. CALIOP

4.2.1 Cloud top automatic detection

For each L1 532 nm version 4.10 CALIOP backscatter product collocated with RO profiles, an automatic cloud top detection was performed using a customized Matlab algorithm. The collocation thresholds were kept at ±0.2° and ±12h between RO profiles and CALIOP backscatter products to provide cloud top altitudes consistent between the instruments. The first step of the cloud top detection procedure consists in cropping the CALIOP backscatter image according to the collocated RO profile position at +/-14° in latitude and +/-80° in longitude. The objective of this first step is to remove unnecessary information from the CALIOP image and to focus the processing on the zone of interest in order to save computational time. These latitude and longitude ranges are based on a series of tests performed on backscatter images and correspond to the best compromise between image size reduction and loss of volcanic cloud information. Threshold backscatter values are then implemented to remove the noise outside of the range 3×10^{-2} - 7×10^{-4} km^{-1}.sr^{-1} to which volcanic clouds correspond. One median filter (4x3 pixels) and two Wiener filters (4x3 and 2x2 pixels) are then successively applied to the resulting backscatter image to reduce the noise within the threshold range. Below 10 km altitude, the RO bending angle anomaly is quite noisy due to the presence of moisture. We thus focus only on volcanic clouds with a top altitude above 10 km and removed image information below this altitude. The next step of CALIOP data processing is to identify remaining groups of pixels (or clusters) within the image. The Matlab connected components finder is set in this customized algorithm to keep only the clusters combining more than 300 pixels. Amongst these selected clusters the nine biggest ones are kept for the final processing stage. This final stage consists of distinguishing clusters corresponding to volcanic features from the ones corresponding to meteorological clouds. To do so, the aspect ratio of a series of volcanic and meteorological clouds were measured and a threshold value of 0.09 was set as the higher limit for volcanic clouds. Finally, the remaining clusters' top altitudes are measured and an average value calculated and saved in the archive as an estimate of the volcanic cloud top altitude.

4.2.2 Cloud type detection

The cloud type detection was performed using the L2 Vertical Feature Mask (VFM) CALIOP product. These VFM products were collocated at ±0.2° and ±1h with AIRS, IASI and GOME-2 for the purpose of such data is to confirm the presence of certain aerosols types simultaneously with SO_2. Among the different cloud types available in VFM products the types 2, 6, 9 and 10 were of particular interest for this archive since they respectively correspond to dust, elevated smoke, volcanic ash and sulfate/other. For each CALIOP file of the archive, the aerosol subtype was extracted using a customized Matlab routine. This Matlab algorithm reads the VFM data, detects the matching latitude/longitude points of all the CALIOP track collocated with AIRS, IASI and GOME-2 and subsets the latitude and longitude array data based on the chosen spatial window of 2°. The algorithm then extracts the feature sub-type of interest as a function

of the altitude. The final output is subdivided in three levels of altitude -0.5 to 8.2 km, 8.2 to 20.2km and 20.2 to 30.1km.

4.3. Uncertainties

This archive combines five different approaches of volcanic cloud detection and each type of measurement/instrument has its own uncertainties. For AIRS, IASI and GOME-2 the uncertainty depends on many parameters (e.g. thickness of the volcanic cloud, amount of aerosols), one of the most important is the unknown volcanic cloud altitude. Thus, the error is case dependent and a general value of measurement uncertainty cannot be provided. Furthermore, the measurement noise of instruments increases over time due to instrument degradation (Lang et al., 2009; Dikty et al., 2010). However, error budgets of AIRS and IASI can be respectively found in the studies by Prata and Bernardo (2007) and Clarisse et al. (2012), while an uncertainty analysis of GOME-2 is provided by Rix et al. 2012 in the case of the 2010 Eyjafjallajökull's volcanic eruption.

5. Data cross-comparison

Over the past decades, satellite data have proven efficient in volcanic clouds detection through a variety of techniques. Those data are essential in the study of the spreading of such clouds on a global scale but are scattered between the different agencies in charge to process them. This archive gathers satellite data covering 11 VEI 4 eruptions from 2008 to 2016 for a total of 223 days of data coverage (Table 3).

This archive is organized in different sections (Figure 1) with each instrument estimation separated from each other. Several parameters are measured using different instruments, such as SO_2 VCD and cloud top altitude, to allow cross-correlation between the different retrieval algorithms. The database allows the quick visualization of AIRS, IASI, GOME-2, CALIOP and RO data at a given date and time as well as the collocation of any instrument data points with another one. In order to illustrate the use of this archive, we extracted two test cases (Figure 2). The first case (Figure 2a) is the 2008 eruption of Kasatochi volcano for which we selected the 9th of August as reference. The second case (Figure 2b) is the 2009 Sarychev Peak eruption for which we selected the 12th of June as reference. In both cases we considered SO_2 values larger than 3 DU from AIRS, IASI and GOME-2 for 24 hours, RO profiles collocated within ±0.2° and ±12h and CALIOP tracks collocated within ±12h.
In the case of Kasatochi, we selected 4178 AIRS, 1241 IASI and 56 GOME-2 data points with SO_2 VCD larger than 3 DU, 379 CALIOP profiles from 7 different tracks (Figure 2a, blue circles) within ±12h and 100 RO profiles (Figure 2a, red circles) collocated within ±0.2° and ±12h. In the case of Sarychev, 1070 AIRS, 209 IASI and 41 GOME-2 data points, 261 CALIOP profiles from 3 different tracks and 54 RO profiles have been selected with the same criteria. Due to the modular archive structure reported in Figure 1, the user can easily select different time frames, different SO_2 thresholds and collocation period range to be adapted to any purpose.

We have manually verified the correct functioning of the algorithm which collocates the different instruments. We randomly selected several days from different eruptions and compared date, time and coordinates of the acquisitions, then compared the results with those ones automatically provided by the algorithm. We have additionally used a visual validation method for all the samples plotting the SO_2 cloud superimposed to the CALIOP tracks and the RO tangent points. As for the cloud top height, we have collocated the RO and CALIOP estimations with the closest IASI pixel and compared the corresponding values. In Table 4 we report the number of collocations per pairs of instruments with the averaged difference between the estimation. Depending on the eruption the different techniques can give variable performances, for example, the estimations of RO and CALIOP for the Eyjafjallajokull, Kasatochi and Grimsvotn were very close (average difference of 0.3 km, 0.9 km, 1.3 km respectively) while they were large for Calbuco (4.2 km). In general, the cloud top height estimation for eruptions with a large number of colocations (Calbuco, Kasatochi, Nabro and Sarychev Peak) are consistent within the techniques.

6.   Data availability

The raw CALIOP data can be found at https://urs.earthdata.nasa.gov/.

The archive consists of daily files and "eruption" files. For each eruption, we provide access to single daily files or to one file for the whole eruptive period. The files of any eruption are compressed (.zip) NetCDF-4 format (including the daily and whole eruptive period) together with two pdfs (Supplement) describing the file structure. The file names are self-explanatory with daily files following the format *volcanoname_year_month_day.nc* and the eruption files following the format volcanoname.nc. As an example, a user who wishes to access the data corresponding to Kasatochi volcano on 11th of August 2008 will have to look for the file *Kasatochi_2008_08_11.nc*. In case the user wants all the available data for the Kasatochi eruption, they will have to look for the file Kasatochi.nc. The data structure of daily files and volcano files is reported in the Supplement. The data described here will be published with a DOI after final acceptance of this manuscript (Tournigand et al., 2020, http://doi.org/10.5880/fidgeo.2020.016). During the discussion period, the data are accessible via this temporary link: http://pmd.gfz-potsdam.de/panmetaworks/review/0f85d699707efcdc567765bd0dafaaadf94b6df5a531f310167f7e974ea803bf.

7.   Summary and conclusions

This paper presents the first comprehensive archive with information on large SO_2 volcanic clouds since 2006. We collected three different datasets of volcanic SO_2 detection from AIRS, IASI and GOME-2 instruments and co-located the detected pixels with the CALIOP and the GNSS RO products to get information about the cloud vertical structure. The archive provides information about the SO_2 detection (with 3 different algorithms), the cloud top height (with 3 different algorithms), the cloud aerosol type (CALIOP vertical mask feature reference), the atmospheric parameters (bending angle, refractivity, temperature, pressure and specific humidity) and the atmospheric change due to the presence of the volcanic cloud (bending angle anomaly). Up to date, there are no public archives of volcanic clouds which can be used as a reference for further studies and all the information is scattered in different locations and available under different conditions. The aim of this archive and this paper is to provide the users with a complete set of state-of-the-art data. The interest in volcanic clouds detection and monitoring is high and there are still some challenges like the accurate determination of the cloud top height and cloud density to be faced. This archive will make available to the scientific community a relevant number of cases to develop and test new algorithms contributing to improving the accuracy on the estimation of fundamental volcanic clouds parameters. The modular structure of the archive can be easily extended in the future to smaller eruptions (VEI<4) and to other SO_2 estimations, facilitating the inter/cross-comparison between algorithms, allowing to reconstruct the cloud structure and dynamical characteristics and supporting the development of cloud dispersion models.

Author contributions. PYT and VC collected the data. LC provided the IASI dataset. HB provided the GOME-2 dataset. FP provided the AIRS dataset. GK and AS provided the WEGC GNSS RO dataset. PYT, VC and EL conceived the algorithm approach and wrote the code. MH elaborated the CALIOP VFM data. PYT, VC and RB structured and wrote the manuscript. RB conceived the idea, coordinated the team, supervised the project and acquired the funding. EL, MH, LC, HB, FP, GK and AS reviewed the manuscript.

Acknowledgement. The work is accomplished in the frame of the VESUVIO project funded by the Supporting Talent in ReSearch (STARS) grant at Università degli Studi di Padova, IT. We would like to thank Armin Leuprecht for the support and all the suggestions to make the archive technically correct.

References

Global Volcanism Program accessible at https://volcano.si.edu/

The 2008 Eruptions of Okmok and Kasatochi Volcanoes, Alaska, Journal of Geophysical Research: Atmospheres, Special issue, 7 February 2018.

Angerer, B., Ladstädter, F., Scherllin-Pirscher, B., Schwärz, M., Steiner, A.K., Foelsche, U., Kirchengast, G.: Quality aspects of the Wegener Center multi-satellite GPS radio occultation record OPSv5.6, in: Atmospheric Measurement Techniques , 10, 4845–4863, https://doi.org/10.5194/amt-10-4845-2017, 2017.

390 Anthes, R. A., Bernhardt, P. A., Chen, Y., Cucurull, L., Dymond, K. F., Ector, D., Healy, S. B., Ho, S., Hunt, D. C., Kuo, Y., Liu, H., Manning, K., McCormick, C., Meehan, T.K., Randel, W. J., Rocken, C., Schreiner, W. S., Sokolovskiy, S. V., Syndergaard, S., Thompson, D. C., Trenberth, K. E., Wee, T., Yen, N. L., Zeng, Z.: The COSMIC/Formosat/3 mission: early results, in: Bull. Am. Meteorol. Soc., 89, 313–333. http://dx.doi.org/10.1175/BAMS-89-3-313, 2008.

395 Beyerle, G., Schmidt, T., Michalak, G., Heise, S., Wickert, J., Reigber, C.: GPS radio occultation with GRACE: atmospheric profiling utilizing the zero difference technique, in: Geophys. Res. Lett., 32, L13806. http://dx.doi.org/10.1029/2005GL023109, 2005.

Bignami, C., Corradini, S., Merucci, L., de Michele, M., Raucoules, D., De Astis, G., ... & Piedra, J: Multisensor satellite monitoring of the 2011 Puyehue-Cordon Caulle eruption, in: IEEE journal of selected topics in applied earth
400 observations and remote sensing, 7, 2786-2796, https://doi.org/10.1109/JSTARS.2014.2320638, 2014.

Biondi, R., Steiner, A. K., Kirchengast, G., Brenot, H., & Rieckh, T.: Supporting the detection and monitoring of VCs: A promising new application of Global Navigation Satellite System radio occultation, in: Adv. Space Res. , 60, 2707–2722, https://doi.org/10.1016/j.asr.2017.06.039, 2017.

Bourassa, A. E., Robock, A., Randel, W. J., Deshler, T., Rieger, L. A., Lloyd, N. D., ... & Degenstein, D. A: Large
405 volcanic aerosol load in the stratosphere linked to Asian monsoon transport, in: Science, 337, 78-81, https://doi.org/ 10.1126/science.1219371, 2012.

Brenot, H., Theys, N., Clarisse, L., van Geffen, J., van Gent, J., Van Roozendael, M., et al.: Support to Aviation Control Service (SACS): an online service for near-real-time satellite monitoring of volcanic plumes, in: Natural Hazards and Earth System Sciences, 14(5), 1099–1123. https://doi.org/10.5194/nhess-14-1099-2014, 2014.

410 Brown, S. K., Crosweller, H. S., Sparks, R. S. J., Cottrell, E., Deligne, N. I., Guerrero, N. O., ... Takarada, S.: Characterisation of the Quaternary eruption record: Analysis of the Large Magnitude Explosive Volcanic Eruptions (LaMEVE) database, in: Journal of Applied Volcanology, 3(1), 1–22. https://doi.org/10.1186/2191-5040-3-5, 2014.

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., Balis, D.: The vertical distribution of volcanic SO2 plumes measured by IASI, in: Atmospheric
415 Chemistry and Physics, 16(7), 4343–4367. https://doi.org/10.5194/acp-16-4343-2016, 2016.

Carn, S. A., & Lopez, T. M.: Opportunistic validation of sulfur dioxide in the Sarychev Peak volcanic eruption cloud, in: Atmospheric Measurement Techniques, 4(9), 1705, https://doi.org/10.5194/amt-4-1705-2011, 2011.

Carn, S. A., Clarisse, L., & Prata, A. J.: Multi-decadal satellite measurements of global volcanic degassing, in: Journal of Volcanology and Geothermal Research, 311, 99–134. https://doi.org/10.1016/j.jvolgeores.2016.01.002, 2016.

420 Carn, S. A., Fioletov, V. E., Mclinden, C. A., Li, C., & Krotkov, N. A.: A decade of global volcanic SO2 emissions measured from space, in: Scientific Reports, 7, 1–12. https://doi.org/10.1038/srep44095, 2017.

Carn, S. A.: Multi-Satellite Volcanic Sulfur Dioxide L4 Long-Term Global Database V3, Greenbelt, MD, USA, Goddard Earth Science Data and Information Services Center (GES DISC), 10.5067/MEASURES/SO2/DATA404, 2019.

425 Cigala, V., Biondi, R., Prata, A. J., Steiner, A. K., Kirchengast, G., & Brenot, H.: GNSS Radio Occultation Advances the Monitoring of Volcanic Clouds: The Case of the 2008 Kasatochi Eruption, in: Remote Sensing, 11, 2199, https://doi.org/10.3390/rs11192199, 2019.

Clarisse, L., Hurtmans, D., Clerbaux, C., Hadji-Lazaro, J., Ngadi, Y., and Coheur, P.-F.: Retrieval of sulphur dioxide from the infrared atmospheric sounding interferometer (IASI), in: Atmos. Meas. Tech., 5, 581–594,
430   https://doi.org/10.5194/amt-5-581-2012, 2012.

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

Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi,
435   A., Turquety, S., Wespes, C., Coheur, P.-F.: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder, in: Atmos. Chem. Phys., 9, 6041-6054, https://doi.org/10.5194/acp-9-6041-2009, 2009.

Cooke, M. C., Francis, P. N., Millington, S., Saunders, R., & Witham, C.: Detection of the Grímsvötn 2011 volcanic eruption plumes using infrared satellite measurements, in: Atmospheric Science Letters, 15, 321-327, https://doi.org/10.1002/asl2.5062, 014.

440   de Moor, J. M., Kern, C., Avard, G., Muller, C., Aiuppa, A., Saballos, A., … Fischer, T. P.: A New Sulfur and Carbon Degassing Inventory for the Southern Central American Volcanic Arc: The Importance of Accurate Time-Series Data Sets and Possible Tectonic Processes Responsible for Temporal Variations in Arc-Scale Volatile Emissions, in: Geochemistry, Geophysics, Geosystems, 18, 4437–4468. https://doi.org/10.1002/2017GC007141, 2017.

Doeringer, D., Eldering, A., Boone, C. D., González Abad, G., & Bernath, P. F.: Observation of sulfate aerosols and
445   SO2 from the Sarychev volcanic eruption using data from the Atmospheric Chemistry Experiment (ACE), in: Journal of Geophysical Research: Atmospheres, 117, D3, https://doi.org/10.1029/2011JD0165562, 012.

Dikty, S., Richter, A., Bovensmann, H., Wittrock, F., Weber, M., Noël, S., Burrows, J. P., Munro, R., and Lang, R.: GOME-2 level 2 products at IUP Bremen and first results on the quantification of the effects of optical degradation, paper presented at Meteorological Satellite Conference, Eur. Organ. for the Exploit. of Meteorol. Satell., Cordoba,
450   Spain, 20–24 Sept, 2010.

EOPAC Team (2019), GNSS Radio Occultation Record (OPS 5.6 2001–2018), University of Graz, Austria, https://doi.org/10.25364/WEGC/OPS5.6:2019.1

Fee, D., Yokoo, A., & Johnson, J. B.: Introduction to an open community infrasound dataset from the actively erupting Sakurajima Volcano, Japan, in: Seismological Research Letters, 85, 1151–1162. https://doi.org/10.1785/0220140051,
455   2014.

Flemming, J., & Inness, A.: Volcanic sulfur dioxide plume forecasts based on UV satellite retrievals for the 2011 Grímsvötn and the 2010 Eyjafjallajökull eruption, in: Journal of Geophysical Research: Atmospheres, 118, 10-172, https://doi.org/10.1002/jgrd.507532, 013.

Flentje, H., Claude, H., Elste, T., Gilge, S., Köhler, U., Plass-Dülmer, C., ... & Fricke, W.: The Eyjafjallajökull
460   eruption in April 2010-detection of volcanic plume using in-situ measurements, ozone sondes and lidar-ceilometer profiles, in: Atmospheric Chemistry and Physics, 10, 10085, https://doi.org/ 10.5194/acp-10-10085-2010, 2010.

Foelsche, U., Scherllin-Pirscher, B., Ladstädter, F., Steiner, A.K., Kirchengast, G.: Refractivity and temperature climate records from multiple radio occultation satellites consistent within 0.05%, in: Atmos. Meas. Tech., 4, 2007–2018. http://dx.doi.org/10.5194/amt-4- 2007-2011, 2011.

465   Fromm, M., Kablick III, G., Nedoluha, G., Carboni, E., Grainger, R., Campbell, J., & Lewis, J.: Correcting the record of volcanic stratospheric aerosol impact: Nabro and Sarychev Peak. Journal of Geophysical Research: Atmospheres, 119, 10-343, https://doi.org/10.1002/2014JD0215072, 2014.

Ge, C., Wang, J., Carn, S., Yang, K., Ginoux, P., & Krotkov, N.: Satellite-based global volcanic SO2 emissions and sulfate direct radiative forcing during 2005-2012, in: Journal of Geophysical Research: Atmospheres, 121, 3446–
470   3464. https://doi.org/10.1002/2015JD023134, 2016.

Griessbach, S., Hoffmann, L., Spang, R., & Riese, M.: Volcanic ash detection with infrared limb sounding: MIPAS observations and radiative transfer simulations, in: Atmos. Meas. Tech, 7, 1487-1507, https://doi.org/ 10.5194/amt-7-1487-2014, 2014.

475   Hajj, G. A., Ao, C. O., Iijima, B. A., Kuang, D., Kursinski, E. R., Mannucci, A.J., Meehan, T. K., Romans, L. J., de la Torre Juarez, M., Yunck, T. P.: CHAMP and SAC-C atmospheric occultation results and intercomparisons, in: J. Geophys. Res., 109, D06109. http://dx.doi.org/10.1029/2003JD003909, 2004.

Houghton, B. F., Swanson, D. A., Rausch, J., Carey, R. J., Fagents, S. A., & Orr, T. R.: Pushing the volcanic explosivity index to its limit and beyond: Constraints from exceptionally weak explosive eruptions at Kilauea in 2008, in: Geology, 41, 627–630. https://doi.org/10.1130/G34146.1, 2013.

480   Kravitz, B., Robock, A., Bourassa, A., Deshler, T., Wu, D., Mattis, I., ... & Duck, T. J.: Simulation and observations of stratospheric aerosols from the 2009 Sarychev volcanic eruption, in: Journal of Geophysical Research: Atmospheres, 116, D18, https://doi.org/10.1029/2010JD015501, 2011.

Kristiansen, N. I., Prata, A. J., Stohl, A., Carn S. A.: Stratospheric volcanic ash emissions from the 13 February 2014 Kelut eruption, in Geophysical Research Letters, 42, 588-596, https://doi.org/10.1002/2014GL062307, 2015.

485   Kursinski, E. R., Hajj, G. A., Schofield, J. T., Linfield, R. P.: Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System, in: J. Geophys. Res., 102(D19), 23,423-429,465. https://doi.org/10.1029/97JD01569, 1997.

Lang, R., Munro, R., Livschitz, Y., Dyer, R., and Lacan, A.: GOME-2 FM3 long-term in-orbit degradation—Basic signatures after 2nd throughput test, Tech Rep. EUM.OPS-EPS.DOC.09.0464, Eur. Organ. for the Exploit. of
490   Meteorol. Satell., Darmstadt, Germany, 2009.

JS Lopes, F., Silva, J. J., Antuña Marrero, J. C., Taha, G., & Landulfo, E.: Synergetic aerosol layer observation after the 2015 Calbuco volcanic eruption event, in: Remote Sensing, 11, 195, https://doi.org/10.3390/rs11020195, 2019.

Luntama, J. P., Kirchengast, G., Borsche, M., Foelsche, U., Steiner, A., Healy, S., von Engeln, A., O'Clerigh, E., Marquardt, C.: Prospects of the EPS GRAS mission for operational atmospheric applications, in: Bull. Am. Meteorol.
495   Soc., 89(12), 18631875. https://doi.org/10.1175/2008BAMS2399.1, 2008.

Marenco, F., Johnson, B., Turnbull, K., Newman, S., Haywood, J., Webster, H., & Ricketts, H.: Airborne lidar observations of the 2010 Eyjafjallajökull volcanic ash plume, in: Journal of Geophysical Research: Atmospheres, 116, D20, https://doi.org/10.1029/2011JD016396, 2011.

Marzano, F. S., Lamantea, M., Montopoli, M., Herzog, M., Graf, H., & Cimini, D.: Microwave remote sensing of the
500   2011 Plinian eruption of the Grímsvötn Icelandic volcano, in: Remote sensing of Environment, 129, 168-184, https://doi.org/10.1016/j.rse.2012.11.005, 2013.

Marzano, F. S., Corradini, S., Mereu, L., Kylling, A., Montopoli, M., Cimini, D., ... & Stelitano, D.: Multisatellite Multisensor Observations of a Sub-Plinian Volcanic Eruption: The 2015 Calbuco Explosive Event in Chile, in: IEEE Transactions on Geoscience and Remote Sensing, 56, 2597-2612, https://doi.org/ 10.1109/TGRS.2017.2769003,
505   2018.

Munro, R., Eisinger, M., Anderson, C., Callies, J., Corpaccioli, E., Lang, R., Lefebvre, A., Livschitz, Y., Pérez Albiñana, A.: GOME-2 on MetOp. In: Proc. of The 2006 EUMETSAT Meteorological Satellite Conference, Helsinki, Finland, p. 48, 2006.

Newhall, C. G., & Self, S.: The volcanic explosivity index (VEI) an estimate of explosive magnitude for historical
510   volcanism, in: Journal of Geophysical Research, 87(C2), 1231. https://doi.org/10.1029/JC087iC02p01231, 1982.

Pardini, F., Burton, M., Arzilli, F., La Spina, G., & Polacci, M.: SO2 emissions, plume heights and magmatic processes inferred from satellite data: The 2015 Calbuco eruptions, in: Journal of Volcanology and Geothermal Research, 361, 12–24. https://doi.org/10.1016/j.jvolgeores.2018.08.001, 2018.

515 Picquout, A., Lavigne, F., Mei, E. T. W., Grancher, D., Noer, C., Vidal, C. M., & Hadmoko, D. S.: Air traffic disturbance due to the 2010 Merapi volcano eruption, in: Journal of volcanology and geothermal research, 261, 366-375, https://doi.org/10.1016/j.jvolgeores.2013.04.005, 2013.

Platt, U. and Stutz, J.: Differential optical absorption spectroscopy: Principles and Applications., Springer-Verlag Berlin Heidelberg, 2008.

520 Prata, A. J., & Bernardo, C.: Retrieval of volcanic SO2 column abundance from Atmospheric Infrared Sounder data, in: Journal of Geophysical Research: Atmospheres, 112(D20). https://doi.org/10.1029/2006JD007955, 2007.

Prata, A. J., & Prata, A. T.: Eyjafjallajökull volcanic ash concentrations determined using Spin Enhanced Visible and Infrared Imager measurements, in: Journal of Geophysical Research: Atmospheres, 117, D20, https://doi.org/10.1029/2011JD016800, 2012.

525 Prata, A. J., Gangale, G., Clarisse, L., & Karagulian, F.: Ash and sulfur dioxide in the 2008 eruptions of Okmok and Kasatochi: Insights from high spectral resolution satellite measurements, in: Journal of Geophysical Research, 115(22), D00L18. https://doi.org/10.1029/2009JD013556, 2010.

Prata, F., Woodhouse, M., Huppert, H. E., Prata, A., Thordarson, T., & Carn, S.: Atmospheric processes affecting the separation of volcanic ash and SO2 in volcanic eruptions: Inferences from the May 2011 Grímsvötn eruption, in: Atmospheric Chemistry and Physics, 17(17), 10709–10732. https://doi.org/10.5194/acp-17-10709-2017, 2017.

530 Rix, M., Valks, P., Hao, N., Loyola, D. G., Schlager, H., Huntrieser,H. H., Flemming, J., Koehler, U., Schumann, U., and Inness, A.: Volcanic SO2, BrO and plume height estimations using GOME-2 satellite measurements during the eruption of Eyjafjallajökull in May 2010, in: J. Geophys. Res., 117, D00U19, https://doi.org/10.1029/2011JD016718, 2012.

Robock, A.: Climatic impacts of volcanic eruptions, in: The encyclopedia of volcanoes (pp. 935-942). Academic 535 Press, 2015.

Rybin, A., Chibisova, M., Webley, P., Steensen, T., Izbekov, P., Neal, C., & Realmuto, V.: Satellite and ground observations of the June 2009 eruption of Sarychev Peak volcano, Matua Island, Central Kuriles, in: Bulletin of Volcanology, 73, 1377-1392, https://doi.org/10.1007/s00445-011-0481-0, 2011.

Scherllin-Pirscher, B., Steiner, A.K., Kirchengast, G., Kuo, Y.-H., Foelsche, U.: Empirical analysis and modeling of 540 errors of atmospheric profiles from GPS radio occultation, in: Atmos. Meas. Tech., 4, 1875–1890. http://dx.doi.org/10.5194/amt-4-1875-2011, 2011.

Schnetzler, C. C., Bluth, G. J. S., Krueger, A. J., & Walter, L. S.: A proposed volcanic sulfur dioxide index (VSI), in: Journal of Geophysical Research: Solid Earth, 102, 20087-20091, https://doi.org/10.1029/97JB01142, 1997.

Stohl, A., Prata, A. J., Eckhardt, S., Clarisse, L., Durant, A., Henne, S., ... & Stebel, K.: Determination of time-and 545 height-resolved volcanic ash emissions and their use for quantitative ash dispersion modeling: the 2010 Eyjafjallajökull eruption, in: Atmospheric Chemistry and Physics, 11, 4333-4351. https://doi.org/ 10.5194/acp-11-4333-2011, 2011.

Stocker, M., Ladstädter, F., Wilhelmsen, H., & Steiner, A. K.: Quantifying Stratospheric Temperature Signals and Climate Imprints From Post-2000 Volcanic Eruptions, in: Geophysical Research Letters, 46, 12486-12494, 550 https://doi.org/10.1029/2019GL084396, 2019.

Surono, Jousset, P., Pallister, J., Boichu, M., Buongiorno, M. F., Budisantoso, A., … Lavigne, F.: The 2010 explosive eruption of Java's Merapi volcano—A '100-year' event, in: Journal of Volcanology and Geothermal Research, 241–242, 121–135, https://doi.org/10.1016/j.jvolgeores.2012.06.018, 2012.

Susskind, J., Barnet, C. D., & Blaisdell, J. M.: Retrieval of atmospheric and surface parameters from 555 AIRS/AMSU/HSB data in the presence of clouds, in: IEEE Transactions on Geoscience and Remote Sensing, 41, 390-409, 10.1109/TGRS.2002.808236, 2003.

Telling, J., Flower, V. J. B., & Carn, S. A.: A multi-sensor satellite assessment of SO2 emissions from the 2012–13 eruption of Plosky Tolbachik volcano, Kamchatka, in: Journal of Volcanology and Geothermal Research, 307, 98-106, https://doi.org/10.1016/j.jvolgeores.2015.07.010, 2015.

560   Theys, N., Campion, R., Clarisse, L., Brenot, H., van Gent, J., Dils, B., Corradini, S., Merucci, L., Coheur, P.-F., Van Roozendael, M., Hurtmans, D., Clerbaux, C., Tait, S., and Ferrucci, F.: Volcanic SO2 fluxes derived from satellite data: a survey using OMI, GOME-2, IASI and MODIS, Atmos. Chem. Phys., 13, 5945–5968, https://doi.org/10.5194/acp-13-5945-2013, 2013.

Theys, N., De Smedt, I., Van Roozendael, M., Froidevaux, L., Clarisse, L., & Hendrick, F.: First satellite detection of
565   volcanic OClO after the eruption of Puyehue-Cordón Caulle, in: Geophysical Research Letters, 41, 667-672, https://doi.org/10.1002/2013GL058416, 2014.

Tournigand, P.-Y., Cigala, V., Lasota, E., Hammouti, M., Clarisse, L., Brenot, H., Prata, F., Kirchengast, G., Steiner, A., Biondi, R.: A comprehensive archive of large SO2 volcanic clouds in 2000s. GFZ Data Services. http://doi.org/10.5880/fidgeo.2020.016, 2020.

570   Turner, M. B., Bebbington, M. S., Cronin, S. J., & Stewart, R. B.: Merging eruption datasets: Building an integrated Holocene eruptive record for Mt Taranaki, New Zealand, in: Bulletin of Volcanology, 71(8), 903–918, https://doi.org/10.1007/s00445-009-0274-x, 2009.

Vernier, J. P., Fairlie, T. D., Deshler, T., Natarajan, M., Knepp, T., Foster, K., ... & Trepte, C.: In situ and space-based observations of the Kelud volcanic plume: The persistence of ash in the lower stratosphere, in: Journal of Geophysical
575   Research: Atmospheres, 121, 11-104, https://doi.org/10.1002/2016JD025344, 2016.

Wickert, J., Reigber, C., Beyerle, G., König, R., Marquardt, C., Schmidt, T., Grunwaldt, L., Galas, R., Meehan, T. K., Melbourne, W. G., Hocke, K.: Atmosphere sounding by GPS radio occultation: First results from CHAMP, in: Geophysical research letters, 28(17), 3263-3266, https://doi.org/10.1029/2001GL013117, 2001.

Williams-Jones, G., & Rymer, H.: Hazards of volcanic gases, in: The Encyclopedia of Volcanoes (pp. 985-992),
580   Academic Press, 2015.

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., Young, S. A.: Overview of the CALIPSO mission and CALIOP data processing algorithms, in: J. Atmos. Oceanic Technol., 26, 2310–2323. http://dx.doi.org/10.1175/2009JTECHA1281.1, 2009.

Yang, K., Liu, X., Bhartia, P. K., Krotkov, N., Carn, S., Hughes, E., Krueger, A., Spurr, R., and Trahan, S.: Direct
585   retrieval of sulfur dioxide amount and altitude from spaceborne hyperspectral UV measurements: Theory and application, in: J. Geophys. Res., 115, D00L09, https://doi.org/10.1029/2010JD013982, 2010.

Zeng, Z., Sokolovskiy, S., Schreiner, W. S., Hunt, D.: Representation of Vertical Atmospheric Structures by Radio Occultation Observations in the Upper Troposphere and Lower Stratosphere: Comparison to High-Resolution Radiosonde Profiles, in: J. Atmos. Ocean. Technol., 36(4), 655–670. http://dx.doi.org/10.1175/JTECH-D-18-0105.1,
590   2019.

Table 1. Summary of the volcanoes and related eruption selected for the database. The following information is provided: the name of each volcano; the eruption start date as provided by the GVP; the spatial location of the volcano in latitude and longitude; the plume/cloud height as a range estimated from IASI, CALIOP and GNSS RO (see details below) and the SO₂ mass loading in Tg as reported in the literature.

Volcano name	Eruption start date	Location Latitude/Longitude (degree)	Cloud average height (km) - Sensor	SO₂ mass loading (Tg) (Reference) Sensor
Okmok	12.07.2008	53.397/-168.166	12.6 - IASI	0.12 (Spinei et al., 2010) OMI

			12.0 - CALIOP 12.5 - RO	0.3 (Prata et al. 2010) AIRS 0.09 (Carn et al., 2016) IASI
Kasatochi	07.08.2008	52.172/-175.509	11.7 - IASI 12.0 - CALIOP 12.4 - RO	2.7 (Corradini et al., 2010) MODIS 1.2 (Prata et al., 2010) AIRS 2.2 (Krotkov et al., 2010) OMI 2.0 (Yang et al., 2010) OMI 1.7 (Karagulian et al.) IASI 1.7 (Kristiansen et al., 2010) Multi-sensor 1.6 (Clarisse et al., 2012) IASI 1.6 (Nowlan et al., 2011) GOME-2
Sarychev	15.06.2009	48.092/153.200	12.2 - IASI 12.9 - CALIOP 12.3 - RO	1.2 (Haywood et al., 2010) IASI
Eyjafjallajokull	20.03.2010	63.633/19.633	8.0 - IASI 12.2 - CALIOP 12.3 - RO	0,17 (Boichu et al., 2013) IASI 1.2 (Rix et al., 2012) GOME-2 0.18 (Carboni et al., 2012) IASI 0.06 (Pugnaghi et al., 2016) MODIS
Merapi	04.11.2010	-7.542/110.442	12.4 - IASI 14.5 - CALIOP 16.1 - RO	0.44 (Surono et al., 2012) Multi-sensor
Grimsvotn	21.05.2011	64.416/-17.316	10.8 - IASI 12.2 - CALIOP 11.7 - RO	0.24 (Prata et al., 2017) 0.38 (Carn et al., 2016) 0.4 (Sigmarsson et al., 2013) OMI+IASI 0.61 (Moxnes et al., 2014)
Nabro	13.06.2011	13.370/41.700	12.2 - IASI 14.3 - CALIOP 15.3 - RO	4.5 (Theys et al., 2013) Multi-sensor
Puyehue-Cordon Caulle (PCC)	04.06.2011	-40.590/-72.117	12.2 - IASI NA - CALIOP 12.5 - RO	0.2 (Theys et al., 2013) IASI
Tolbachik	27.11.2012	55.832/160.326	8.9 - IASI 11.4 - CALIOP 11.7 - RO	0.2 (Telling et al. 2015) Multi-sensor 0.09 (Carn et al., 2016) IASI
Kelut	13.02.2014	-7.939/112.307	17.6 - IASI NA - CALIOP 16.9 - RO	0.2 (Carn et al., 2016) OMI 0,19 (Carn et al., 2016) IASI
Calbuco	24.04.2015	-41.328/-72.607	16.4 - IASI 12.4 - CALIOP 14.8 - RO	0.3 (Pardini et al., 2018)

Table 2. Summary of data used to build the archive.

Sensor	Satellite(s)	Vertical resolution	Spatial resolution	Estimation	Algorithm reference
AIRS *(Infrared)*	Aqua (A-Train)	NA	13.5 km	SO₂ VCD	Prata and Bernardo, 2007
IASI *(Infrared)*	MetOp-A/B	NA	12 km	SO₂ VCD Cloud top height	Clarisse et al., 2012 Clarisse et al., 2014
GOME-2 *(Ultraviolet-Visible)*	MetOp-A/B	NA	40x80 km	SO₂ VCD	Rix et al., 2012
CALIOP *(Lidar)*	CALIPSO (A-Train)	60 m below 20 km 180 m above 20 km	1 km 1.667 km	Cloud top height Cloud type	Winker et al., 2009
GNSS RO *(Microwave)*	CHAMP COSMIC C/NOFS SAC-C GRACE-A Met-Op	100 m in the troposphere 600 m in the stratosphere	50 km in the troposphere 200–300 km in the stratosphere	Bending angle Bangle anomaly Refractivity Temperature Pressure Specific Humidity Cloud top height	Angerer et al., 2017 Cigala et al., 2019

Table 3: Number of days, files, granules and profiles covered by the archive for each volcano in alphabetical order.

Volcano	# of days covered	# of CALIOP profiles	# of AIRS granules	# of IASI scanning steps	# of GOME scanning steps	# of RO profiles
Calbuco	29	12495	350	42740	20992	5362
Eyja	17	3569	76	3980	164369	2624
Grimsvotn	28	6268	147	49824	833541	6007
Kasatochi	23	12897	247	103622	650031	17045
Kelut	2	72	1	1313	2575	83
Merapi	17	1053	27	4919	80193	984
Nabro	26	2463	123	59359	638316	7131
Okmok	26	5678	32	2931	737981	13255
PCC	12	0	76	21528	369992	664
Sarychev	36	11563	127	83533	1035931	16522
Tolbachik	7	617	9	5390	22133	449

600   **Table 4. The average difference between cloud top height estimated with pairs of sensors for each volcano. For each pair we report the average difference of all the collocations and the number of collocations. When there are no collocations the value reported is "/".**

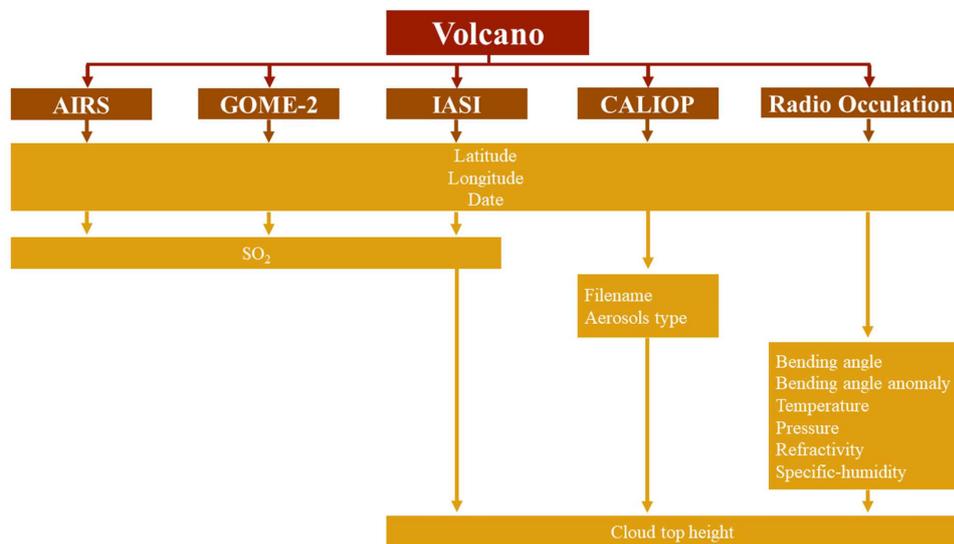
Volcano	RO-CALIOP altitude average (km)	#	RO-IASI altitude average (km)	#	IASI-CALIOP altitude average (km)	#
Calbuco	4.2	39	3.4	867	4.9	308
Eyjafjallajokull	0.3	1	3.7	30	3.5	29
Grimsvotn	0.9	5	3.7	136	2.4	75
Kasatochi	1.3	70	1.2	3855	1.6	997
Kelut	/	0	1.7	20	/	0
Merapi	1.5	1	2.7	127	2.2	70
Nabro	3.4	9	4.3	609	3.6	204
Okmok	3.3	2	1.8	143	2.5	22
PCC	/	0	1.6	193	/	0
Sarychev	1.5	24	1.5	1519	2.8	227
Tolbachik	/	0	3.0	68	/	0

605

Figure 1: Archive schematic organization.

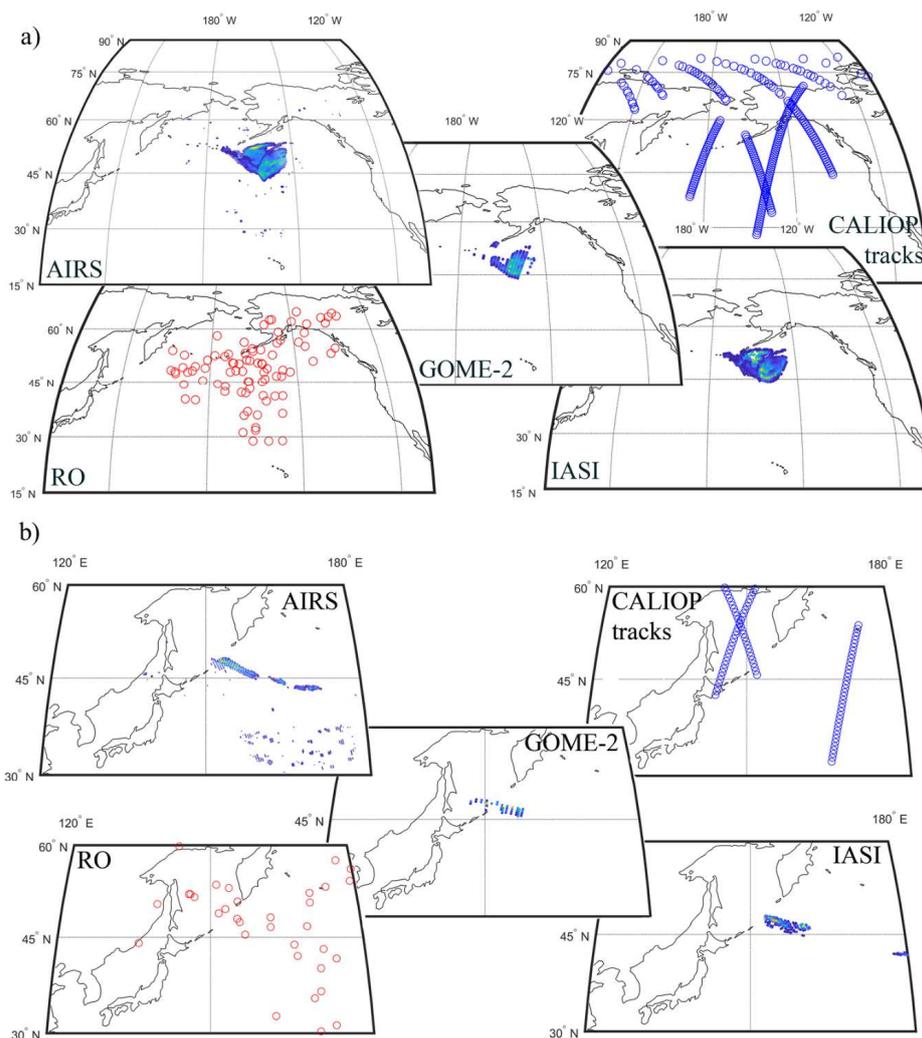



Figure 2: Example of data use and data collocation. Kasatochi cloud on 9th of August 2008 (a) and Sarychev Peak cloud on 12th of June 2009. The central panels show the SO₂ VCD 2-dimensional spreading estimated by AIRS, GOME-2 and IASI, the up-right panel show the CALIOP tracks for which the total attenuated backscatter profile is available. and the bottom-left panel shows the RO profiles collocated with the SO₂.

610