

A machine-learning reference dataset for SO₂ plumes observed by TROPOMI: uncertainties and emission estimates

Douglas P. Finch^{1,2} and Paul I. Palmer^{1,2}

¹School of GeoSciences, University of Edinburgh, Edinburgh, UK

²National Centre for Earth Observation, University of Edinburgh, Edinburgh, UK

Correspondence: Douglas Finch (d.finch@ed.ac.uk)

Abstract. Sulphur dioxide (SO₂) is a major atmospheric pollutant from fossil fuel combustion, metal smelting, and volcanic degassing, impacting human health, acid deposition, and climate forcing. Existing emission inventories are often temporally lagged and spatially coarse, failing to capture high-intensity, sporadic events. To address this, we present a novel, near real-time approach using a U-Net image segmentation model to automatically isolate SO₂ plumes from over 31,000 TROPOMI satellite swaths (Jan 2019–Dec 2024). The model successfully identified 53,993 individual plumes. The highest annual detection rate in 2019 was attributed to massive stratospheric SO₂ injections from the Raikoke and Ulawun volcanic eruptions. Clustering analysis confirmed plume origins around expected volcanic and industrial hotspots (e.g., Popocatepetl, Norilsk), with volcanic sources dominating the top ten clusters. However, the algorithm struggles to detect sources in high-background regions like China, where high SO₂ saturation likely prevents individual plume isolation. This study demonstrates machine learning as a powerful tool for transforming atmospheric monitoring, providing the high-cadence, fine-grained quantification of SO₂ emissions crucial for validating global inventories and ensuring effective environmental management.

1 Introduction

Sulphur dioxide (SO₂) is an atmospheric pollutant predominately produced from fossil fuel combustion for power generation, residential heating, industrial processes (e.g. metal smelting), refineries, shipping, and volcanoes. Within the clean troposphere, the dominant loss of atmospheric SO₂ is oxidation by the hydroxyl radical, resulting in a lifetime of approximately a week although this can vary dramatically depending on environmental conditions. SO₂ contributes to the formation of fine particulate matter that is directly linked with negative health outcomes, particularly cardiovascular diseases (Khalaf et al., 2024). SO₂ also has broader environmental impacts, primarily by forming sulphuric acid via heterogeneous chemistry, which leads to ecosystem damage and building corrosion. The formation of sulfate aerosols affects climate forcing both directly, by scattering incoming sunlight and causing a net cooling of the atmosphere, and indirectly, by perturbing cloud microphysics. In this study, we use machine learning to identify automatically permanent and ephemeral hotspots of SO₂ observed by the TROPOMI satellite instrument and quantify the corresponding emission estimates, carefully curating sources of error. We will focus on large point sources from fossil fuel combustion, copper smelting, and volcanic emissions.

Emissions of SO₂ from fossil fuel combustion hinges on three factors: the sulfur content of the fuel, combustion efficiency, and the deployment of SO₂ scrubber technology. More energy is released during the combustion of coal with a higher number of hydrogen:carbon bonds, which is inversely proportional to the sulfur content. Scrubber technology, widely adopted by coal-fired power plants in developed nations starting in the 1990s (Srivastava et al., 2001), was introduced to meet regulatory requirements to mitigate acid deposition that caused widespread destruction of downwind forest and aquatic ecosystems (Smith, 1872; Driscoll et al., 2001). Global anthropogenic SO₂ emission estimates varied from 93 to 108 Tg yr⁻¹ between 2010 and 2018, with recent years showing lower values (Soulie et al., 2023). While different inventories, such as the Copernicus Atmosphere Monitoring Service (CAMS) and the Emissions Database for Global Atmospheric Research (EDGAR), generally agree within ≈ 4 Tg SO₂ yr⁻¹ estimates diverge in later years, primarily due to differing estimates for power generation ($\approx 44\%$ of CAMS anthropogenic emissions during 2010–2018) and shipping ($\approx 10\%$). International shipping is a significant source due to the high sulfur content of marine fuel, but low-sulfur fuel regulations introduced in 2020 are expected to have significantly reduced this annual emission, consistent with a large reduction in observed ship tracks in cloud perturbations (Watson-Parris et al., 2022).

Extracting copper from mineral ores – predominately chalcopyrite (CuFeS₂) – releases SO₂ to the atmosphere. The smelting process involves injecting mineral particles and oxygen-enriched air into a furnace that is heated ≈ 1500 K, resulting in the sulphide minerals reacting with the injected oxygen that eventually produces SO₂. Currently, most copper refineries are in China, India, Japan, Russia, and Chile, with only a few smaller-capacity plants in the United States and Germany. Developed countries capture the SO₂ waste product but for refineries in developing countries the capture technology may be unaffordable or unavailable.

Volcanoes represent a natural source of SO₂ to the atmosphere. They emit SO₂ in vast quantities during eruptive and during passive degassing periods. Sulphur species is a minor constituent in volcanic magma, compared with water and carbon dioxide. The production and subsequent emission of SO₂ from volcanoes depends on various factors, including the composition and depth of the magma reservoir and the nature of the eruption. Generally, large, explosive (high pressure) volcanic eruptions release more SO₂ to the atmosphere than passive degassing periods, which occur due to the movement of sub-surface magma. Annual volcanic SO₂ emission estimates vary. Estimates inferred from satellite data collected between 2005 and 2015 report an annual mean of 23 ± 2 Tg yr⁻¹ for volcanic degassing, with 30% of those sources exhibiting a positive decadal trend (Carn et al., 2017). However, ground-based data collected at 32 volcanoes over the same period report a mean (median) emission rate of ~ 9.0 (6.8) Tg SO₂ yr⁻¹ (Arellano et al., 2021), and a subset of these ground-based estimates show that some volcanoes degas at a rate too low to be detected by instruments like the Ozone Monitoring Instrument (Carn et al., 2017). Only within the past decade has space-borne sensor technology achieved sufficient sensitivity to accurately detect degassing SO₂ emissions, providing crucial data that complements information gathered by ground-based networks, such as the Network for Observation of Volcanic and Atmospheric Change (Galle et al., 2010).

Traditional bottom-up emission inventories for SO₂ (e.g., EDGAR (EDGARv8.1), CEDS (Hoesly et al., 2018)) suffer from critical limitations for modern monitoring, including significant time lags (often years behind real-time) and poor resolution, providing only mean values over large areas (e.g., 100s of kilometres) and long durations (e.g., monthly). Capturing

sporadic emission events or tracking rapid changes in current sources requires the high temporal resolution offered by near real-time satellite observations, especially emerging data from geostationary instruments (e.g. Global Environmental Monitoring Systems (GEMS, Kim et al. (2020)), Tropospheric Emissions: Monitoring of Pollution (TEMPO Chance et al. (2019)) and Sentinel-4 (Bazalgette Courrèges-Lacoste et al. (2017))) that offer continuous monitoring throughout the sunlit day. To efficiently exploit these massive stream of data, we employ a machine learning model that rapidly highlights SO₂ plumes originating from point sources and an example of how to quickly estimates their emission rates. This capacity for rapid, fine-grained quantification of emission changes is essential for timely intervention and effective atmospheric management. We showcase our approach using data collected by the Tropospheric Monitoring Instrument (TROPOMI) satellite instrument aboard Sentinel-5P.

In the following section, we describe the data used and the development of the model. In section 3 we report our result. We conclude the study in section 4.

2 Data and Methods

2.1 TROPOMI SO₂

We use Level 2 total vertical column density (VCD) SO₂ data retrieved from TROPOMI, a high-resolution UV–Vis–NIR–SWIR spectrometer, onboard the Sentinel-5P satellite, January 2019–December 2024, inclusively. TROPOMI measures the solar radiation backscatter in the UV range (around 310-330 nm) where SO₂ has a distinct absorption feature. This analysis uses the TROPOMI Differential Optical Absorption Spectroscopy (DOAS) retrieval (Theys et al., 2017). Sentinel-5P was launched in October 2017 into a sun-synchronous orbit with a local equatorial overpass time of 13:30. TROPOMI has a swath width of 2600 km divided into 450 across-track pixels, with a pixel resolution of 3.5×5.5 km (across × along track) at nadir for SO₂. This sampling strategy results in near-daily global coverage (Veefkind et al., 2012), subject to cloud-free scenes. In this study, we only use pixels with a quality flag >0.5, as recommended by the TROPOMI Level 2 Product User Manual (Romahn et al.). Because this study uses the full VCD rather than the 1, 7 or 15 km layer products, we acknowledge that some potential SO₂ plumes located above cloud tops may be missed when applying a quality-flag threshold of 0.5. The broader implications of using the full VCD instead of layer-specific products are not assessed here, but this could be incorporated into future model developments and retrieval-selection strategies.

An existing SO₂ detection flag is provided in TROPOMI files as described in the Sentinel-5P/TROPOMI Algorithm Theoretical Basis Document (Theys et al., 2023) built on a detection algorithm Brenot et al. (2014). This detection algorithm combines the SO₂ observation, solar zenith angle, VCD error and proximity to other detections to assign a flag to each pixel. This flag uses five categories: (0) no enhancement, (1) general SO₂ detection, (2) near a known volcano, (3) near a known anthropogenic source, and (4) a potential false positive due to a high solar zenith angle. This study does not include this flagging data in training the model as we want the model to learn SO₂ enhancement and plume shapes without relying on existing methods. The flags indicating if the detection are near a known source (flags 2 and 3) also rely on proximity to a known source, which could adversely affect the ability of model to detect plumes from new sources. We compare the results from our plume detection

model with SO₂ flag provided in section 3. The method presented in this study does not assign a threshold to a potential plume and tests whether a model can learn to correctly identify plumes from the TROPOMI data without pre-defined limits.

2.2 U-Net Image Detection Model

To automatically detect plumes of SO₂ from TROPOMI data, we use a U-Net style fully convolutional network model to perform image segmentation (Ronneberger et al., 2015; Mukhopadhyay et al., 2015). A U-Net model is designed to produce a pixel-by-pixel classification of an image and has been widely used in medical sciences (e.g. tumour detection) as well as in land cover classification in satellite imagery (e.g. Pan et al., 2020; Ulmas and Liiv, 2020; Bokstaller et al., 2021; Filatov and Yar, 2022). It follows the basic principle of convolving an image over successive layers to reduce the spatial dimensions and extract feature information and then using transposed convolutions to rebuild the image to the original shape, the schematic of the model architecture creates a “U” shape, as shown in Figure 1. This figure shows a schematic of the model architecture used for the plume detection algorithm. We add Gaussian noise to the training images to improve the robustness of the model before passing the image through four downsampling blocks. Each block consists of a double convolutional layer, a max-pooling layer and a dropout layer set at 20%. These blocks then feed to one more double convolutional layer to extract patterns in the image before four upsampling blocks create a mask of the plume. Each upsampling block contains a double transpose convolutional layer, a concatenation layer, another dropout layer set at 20% and a double convolutional layer.

This architecture was based upon the original U-Net designs described in Ronneberger et al. (2015) then manually adapted through iterative tuning to fit this study. We found training the model for 20 epochs sufficient for our relatively simple image problem. We found no further model improvement when it was trained over more epochs. As our training dataset was relatively small (1000 images), we chose a batch size of 64 to balance over-fitting with training efficiency. The model uses a sigmoid activation as we wanted the model to make a binary pixel-wise classification of within a plume or not.

Image segmentation models offer significant advantages over traditional image classification (e.g. Finch et al., 2022) because they parse essential information from the background rather than simply assigning a single label to an entire image. For detecting SO₂ plumes, this capability is crucial: it allows for more precise geolocating of the plume and the ability to handle multiple distinct plumes within a single satellite scene. Furthermore, segmentation excels over other feature-parsing methods, such as activation maps (Zhou et al., 2015), because it is trained directly on ground-truth image masks. This direct comparison during training provides a clear, quantifiable performance metric, ensuring reliable results.

Our segmentation model was trained using a custom database of over 1,000 TROPOMI images showing SO₂ plumes, each paired with a precise plume mask manually created by the lead author. These images were sampled randomly from all available TROPOMI files across the full study period and the global domain to minimize the risk of introducing geographical biases. We use the More detail on the creation of the training dataset is provided in Appendix A. To maximize training effectiveness, we augmented this dataset through rotation and flipping, yielding a final training pool of over 4,000 images and corresponding masks. We chose an image size of 32×32 pixels (roughly 112×176 km² at nadir) as it successfully captures most SO₂ plumes. For validation, we trained the model on 80% of the data and tested it on the remaining 20%. The model’s performance was measured using precision (correctness of positive predictions) and recall (completeness of positive detection), yielding scores

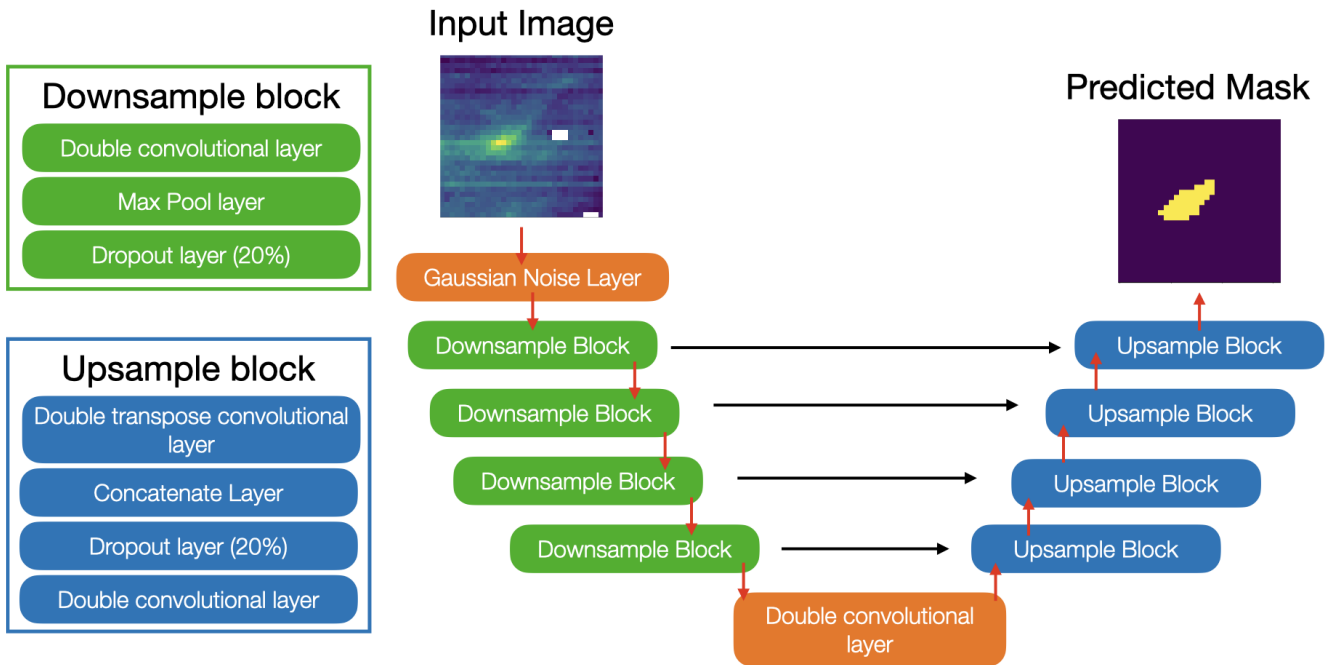


Figure 1. Schematic of the architecture of the U-Net model.

125 of 65.7% and 74%, respectively, giving an F1 score of 0.69. Crucially, the small 32×32 image size disproportionately penalizes minor errors, meaning an offset of just a few pixels drastically lowers the score. Although these scores indicate the performance of the model, the precision and recall percentages are not directly interpretable as plumes missed or not, as segmentation models work on a pixel per pixel basis. For example, the model may correctly detect a plume in the test dataset but draw a smaller or larger mask than the test data and would therefore be penalised. It is also important to note that because the training dataset was
 130 created manually and relied on the authors' judgement of what constitutes a plume, the model cannot outperform the quality of this dataset. Any biases or recurring misclassifications present in the training examples may be learned by the model and subsequently propagated into the final results. An alternative approach would be to augment the training dataset with model-

135 To ensure we capture SO_2 plumes that may be straddling multiple image boundaries, we employ a 32×32 pixel rolling window that moves four pixels at a time across and along the satellite swath. This systematic sampling generates approximately 100,000 images per swath for input into the segmentation model. The model returns each image with a pixel-by-pixel probability of plume presence. We then reconstruct the original swath into an amalgamated mask by taking the median probability of all overlapping pixels. Using the median not only combines the individual predictions but also boosts detection confidence,

140 as an accurately identified plume will appear consistently across multiple overlapping images. We found that a four-pixel step provides adequate coverage to resolve straddling issues while maintaining reasonable computing speed and costs. Crucially, this overlaying method means the final predicted plume shape is not limited by the 32×32 pixel input size, allowing plumes and their corresponding masks to be accurately mapped across the full scale of the swath (which is typically 4172×450 pixels, spanning 2,600 km from pole to pole).

145 To extract the details of individual plumes from the reconstructed satellite swath, we apply connected component analysis (using Open-CV (Bradski, 2000)) to the pixel probability array. This analysis effectively identifies the unique plume masks within the swath and provides the bounding boxes for each one, allowing us to precisely delineate the plume outline. Figure 2 illustrates this capability, showing twenty randomly selected TROPOMI-observed plumes alongside their corresponding predicted plume outlines.

150 Figure 2 demonstrates that the model generally performs well, although some inaccuracies are present. Detecting atmospheric features like SO_2 plumes inherently involves subjectivity, as there is no clear, objective physical boundary for the feature of interest. This human subjectivity is inevitably encoded in the model's training dataset and subsequently reflected in the trained model itself. Continuously refining the model or updating the training dataset is an endless task, so for the practical scope of this paper we have chosen to present results based on a rigorous process involving three training iterations (i.e.,
155 checking model output, expanding the dataset with new examples, and retraining).

We have created a comprehensive database documenting each detected plume, which includes its location, date of detection, plume outline, and an estimate of the emission rate. A bounding box (with a three-pixel buffer) is also recorded, with its limits specifically used to determine the background SO_2 concentration outside the plume. Computationally, processing a single swath is highly efficient, taking only about 15 seconds using a GPU or 60 seconds using CPUs.

160 2.3 Source Emission Rates Estimate

By using the predicted plume outline and modelled wind fields, it is possible to calculate an emission estimate associated with a detected plume. To create a robust estimate of all plumes detected, accurate wind fields would need to be obtained by using the SO_2 layer height from the TROPOMI fields and relating this to a meteorological dataset. A full analysis of the emissions estimates of all detected plumes is computationally expensive and is beyond the scope of this paper but we provide a bare
165 bones example of a plume emission estimate using the following formula:

$$E = \frac{\Delta M \times ws}{L} \times 3600, \quad (1)$$

where ΔM is the mass enhancement of the plume relative to the background in kg, ws is median wind speed in ms^{-1} and L is the length of the plume in metres. The results is then multiplied by 3600 to convert from kg s^{-1} to kg hr^{-1} .

Here we use the 10-metre U and V wind fields from ERA5 reanalysis data, included in most TROPOMI Level 2 files, as a
170 foundational estimate for the wind experienced by the emissions. While these near-surface fields may not perfectly represent transport winds - especially for volcanic SO_2 injected high into the atmosphere - they provide a reasonable starting point.

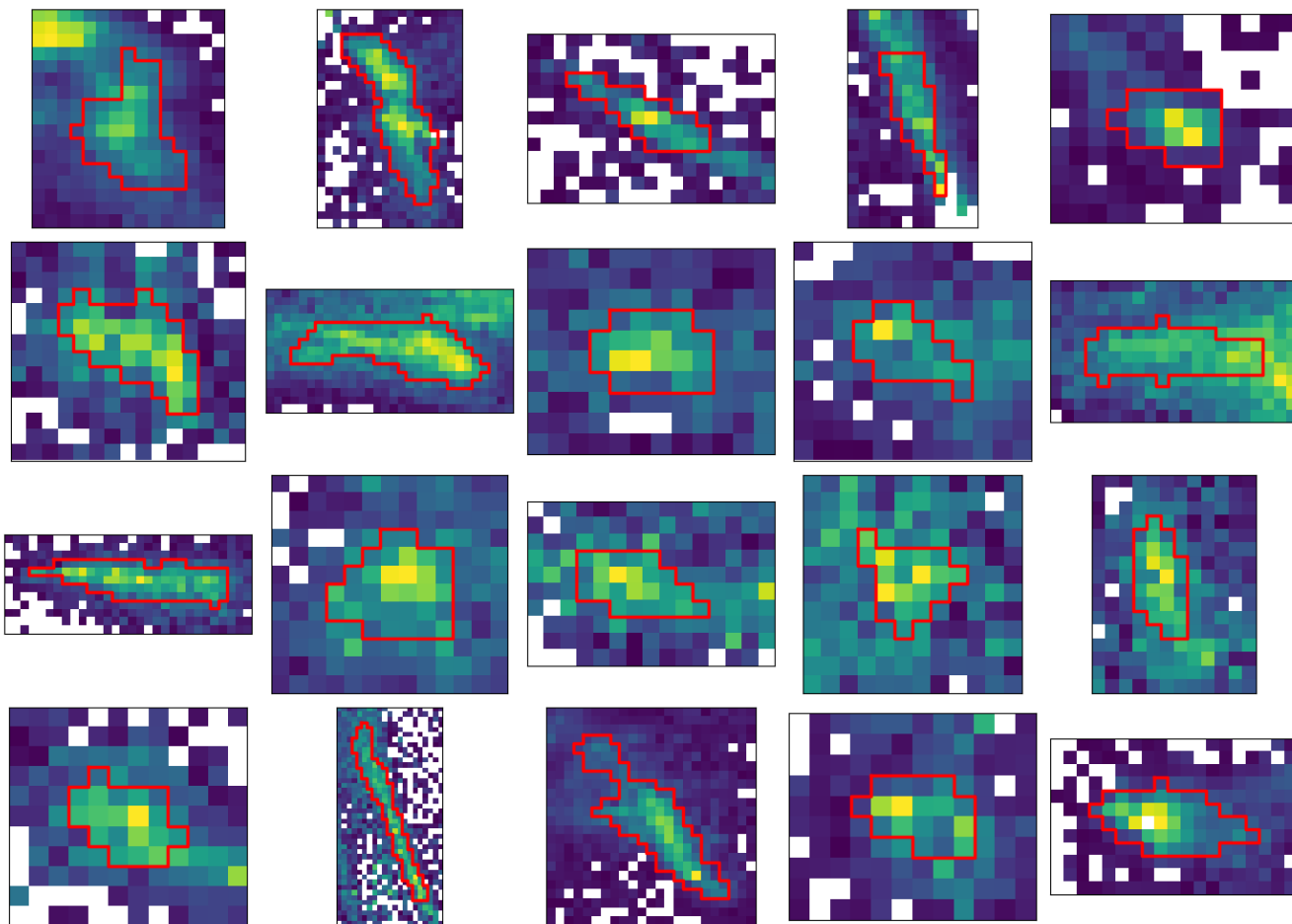


Figure 2. Examples of SO₂ plumes in the TROPOMI data and the predicted plume outline, shown as a red line. Warmer colours denote higher values.

We calculate the area of each pixel within the given scene based on the coordinates of the pixel vertices given in the TROPOMI file which is then used to convert the SO₂ observations from mol m⁻² to grams of SO₂ per pixel. The mass enhancement of the plume is calculated by subtracting the median values of the background, defined as all the pixels within the bounding box but outside the plume boundary.

To estimate the plume length, we fit an ellipse to the plume outline and use the length of its primary axis, trimming it precisely to the plume boundary. Figure 3 visually demonstrates this process, showing the mass enhancement relative to the background median, along with the fitted ellipse and its primary axis. This entire image represents the plume's defined bounding box. The example shown in figure 3 shows a plume detected over the Mount Isa, a copper, zinc, lead and silver mine and refinery in Queensland, Australia (-20.71°N, 139.48°E) on October 17, 2024, with an estimated length of 75.9 km and an emission rate of 36,540 kg hr⁻¹.

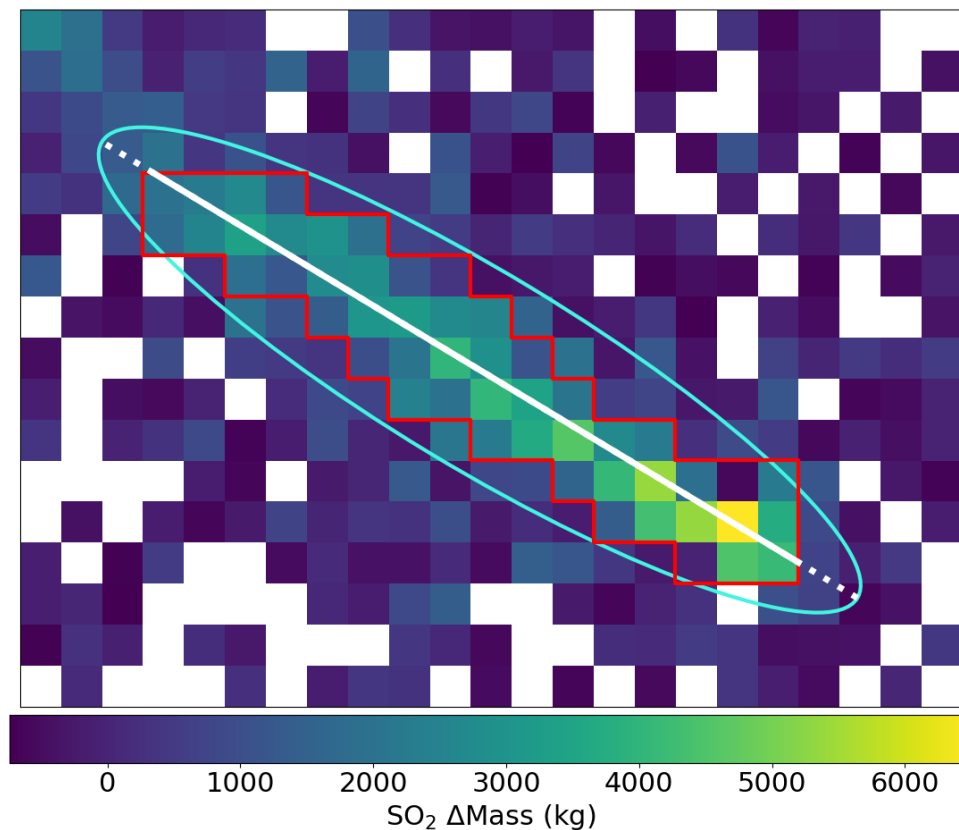


Figure 3. Mass Enhancement (kg) for each plume pixel, shown relative to the background’s median value which is defined by the bounding box area outside the plume boundary. The fitted ellipse is displayed in light blue, with its primary axis shown in white. The dashed line indicates where the primary axis was truncated to estimate the plume length.

3 Results

3.1 Global SO₂ Data

We processed approximately 31,000 Level 2 files from TROPOMI observations spanning January 2019 to December 2024. To ensure the reliability of our output, we implemented a filter requiring each detected plume to contain a minimum of six pixels, as the accuracy of smaller predicted plumes is difficult to validate.

In total, 53,993 plumes were identified over the study period. The year 2019 saw the highest number of detections, exceeding 16,000, while the period from 2020 through 2024 showed more consistent annual counts, ranging from ~6,800 to ~8,000 plumes. To locate the emission source, we employ the coordinates of the maximum SO₂ concentration within the plume

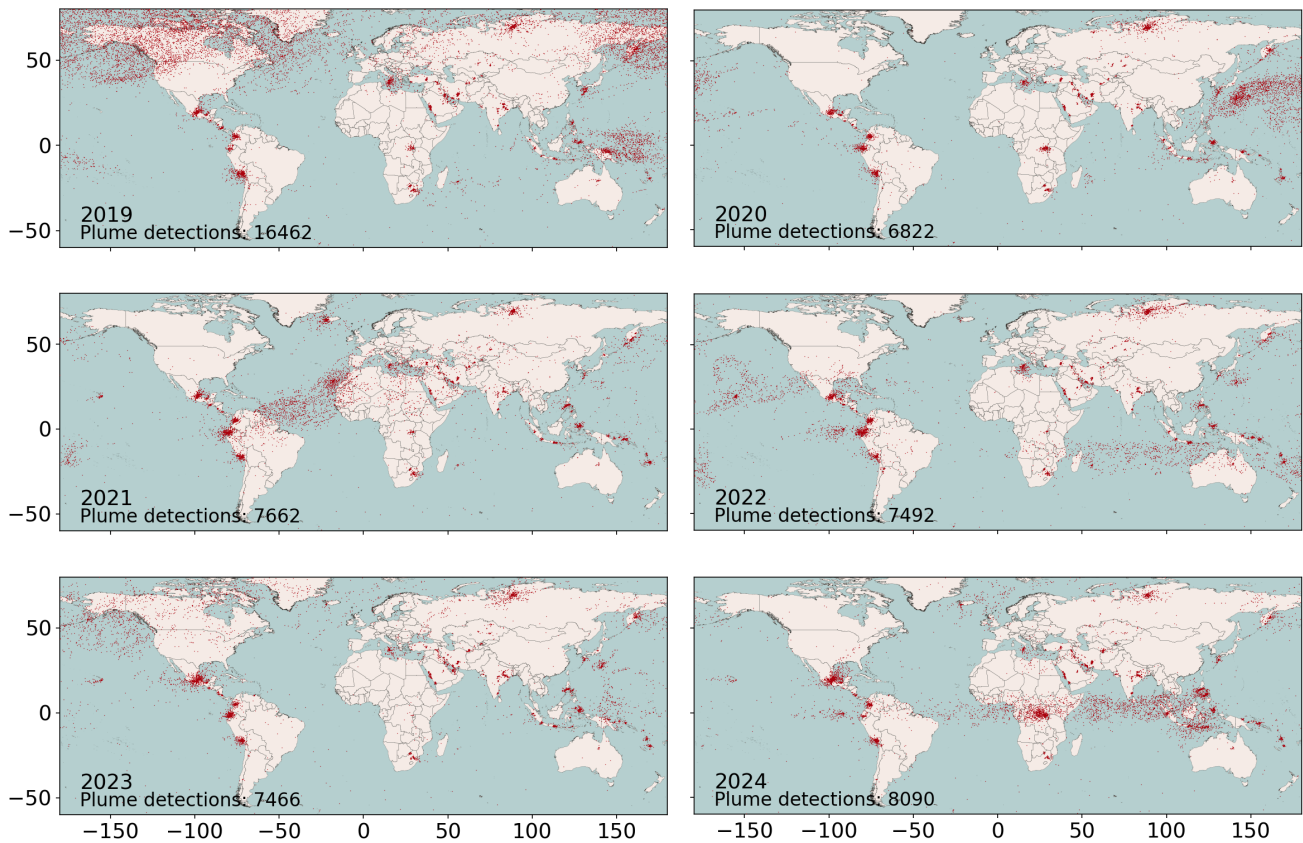


Figure 4. Global, annual locations of the maximum concentration of TROPOMI SO₂ within the predicted plume masks, 2019–2024.

190 boundary as a practical proxy for the origin, noting that this might not perfectly match the true source location. Figure 4 shows
 195 the annual global distribution of these estimated plume origins, 2019–2024.

Initial inspection confirms plumes cluster predictably around known volcanoes and industrial hotspots. However, all years also show regions with a wide, “noisy” spread of detections (e.g., Alaska, Canada, and Eastern Russia annually, plus specific areas like the mid-Atlantic in 2021 or Central Africa in 2024). Figure 5 illustrates the daily plume counts, where each spike
 195 corresponds to these noisy geographical spreads (Figure 4). Critically, all but one spike (Peak I in 2023) align with major volcanic eruptions in the region. Table 1 details the specific volcanic emissions believed responsible for this widespread geographical dispersal. Plume transport away from the source is evident, as detections further from the origin often occur in the days immediately following an eruption event (see Appendix B, Figure B1).

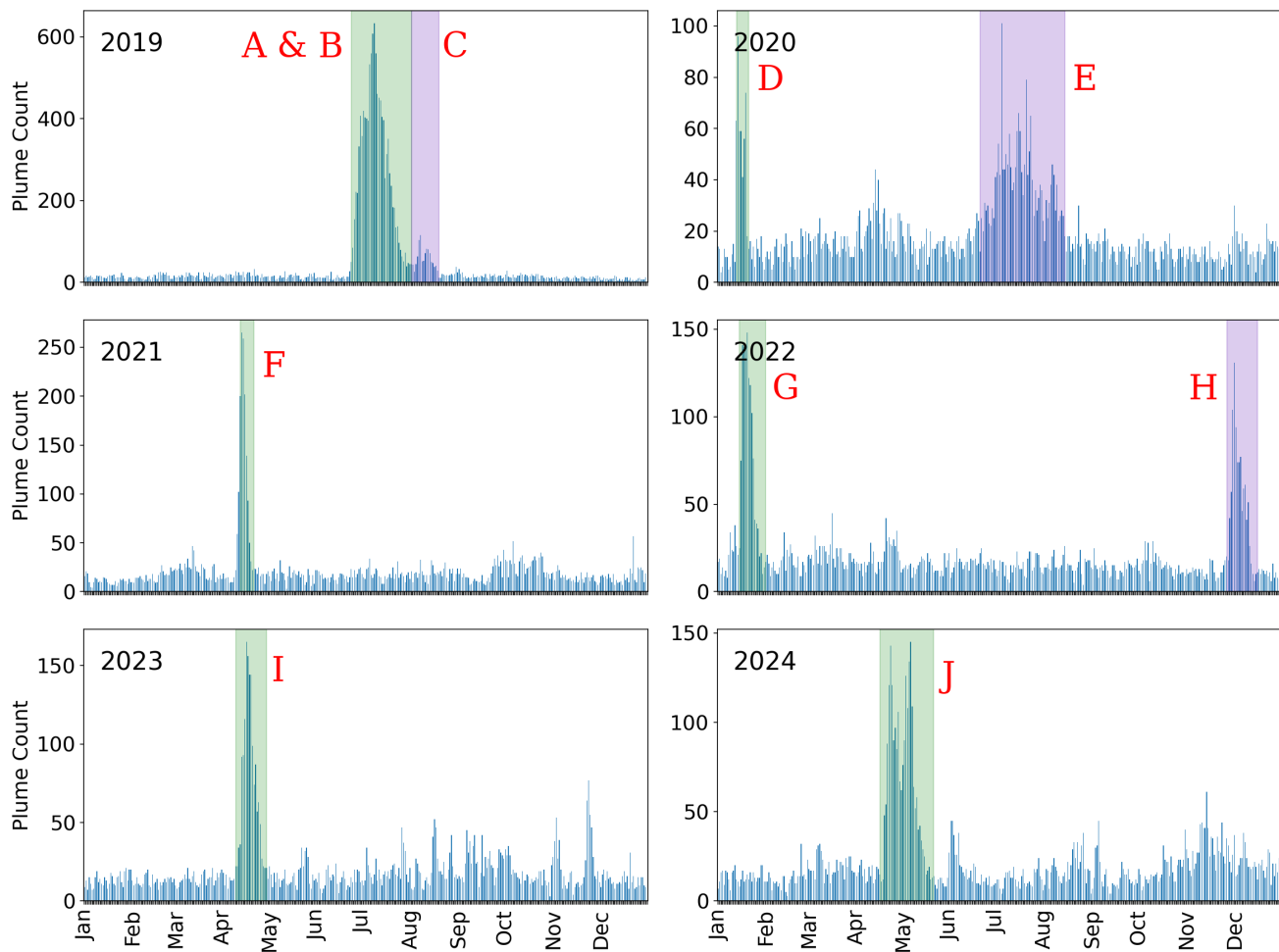


Figure 5. Number of plume detections per day for each year of the study. Highlighted regions show time periods of a higher than usual number of detections. Labels correspond to Table 1.

We attribute the major spike in 2019 plume detections mostly to the coinciding eruptions of Raikoke, Russia, and Ulawun, Papua New Guinea. The Raikoke eruption on June 21st is particularly notable, injecting one of the largest amounts of sulfur dioxide (SO_2) into the stratosphere since the 1991 Mount Pinatubo eruption (Vernier et al., 2024, and references therein).

While Peak I in Figure 5 corresponds to the Shiveluch eruption in Russia on the 11th of April 2023, it also coincides with a surge in detections concentrated over Norilsk in Northern Russia (88.1°W , 69.3°N), as shown in Figure 6. This region is a known major source of SO_2 emissions due to its large metal smelting operations. The precise cause of this specific increase in plume detections, however, remains unknown.

Label	Vocano Name	Country	Coordinates	First Detection Date
A	Raikoke	Russia	153.25°E, 48.29°N	22nd June 2019
B	Ulawun	Papua New Guinea	151.33°E, 5.05°S	26th June 2019
C	Ulawun	Papua New Guinea	151.33°E, 5.05°N	3rd August 2019
D	Taal	Philippines	120.99°E, 14.01°N	12th January 2020
E	Mount Cleveland	United States of America	169.94°W, 52.82°N	18th June 2020
F	La Soufrière	Saint Vincent and the Grenadines	61.18°W, 13.33°N	11th April 2021
G	Hunga Tonga-Hunga Ha'apai	Tonga	175.38°W, 20.55°S	14th January 2022
H	Mauna Loa	United States of America	155.60°W, 19.47°N	26th November 2022
I	Shiveluch	Russia	161.36°W, 56.65°N	11th April 2023
J	Taal	Philippines	120.99°E, 14.01°N	14th April 2024

Table 1. Major volcanic eruptions thought to be responsible for the periods of higher than usual plume detections highlighted and labelled in Figure 5.

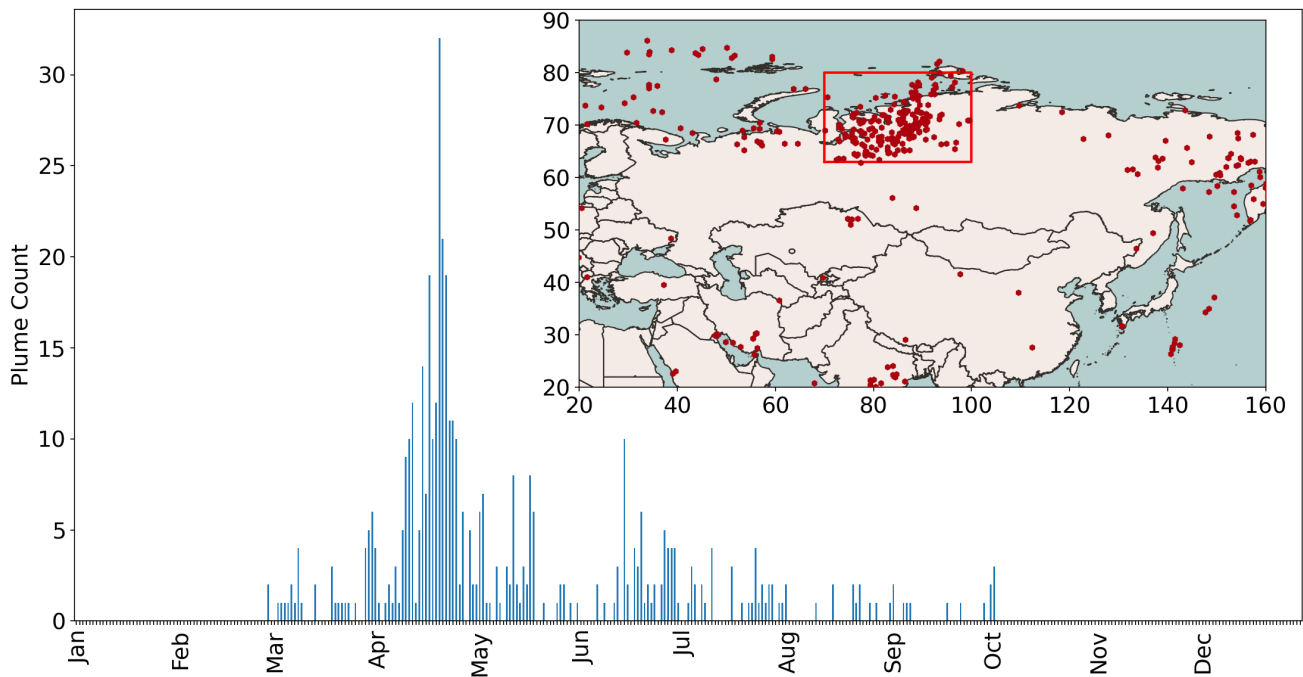


Figure 6. Number of plume detections per day for during 2023 within the red box on the inset map (centred over Norilsk, Russia).

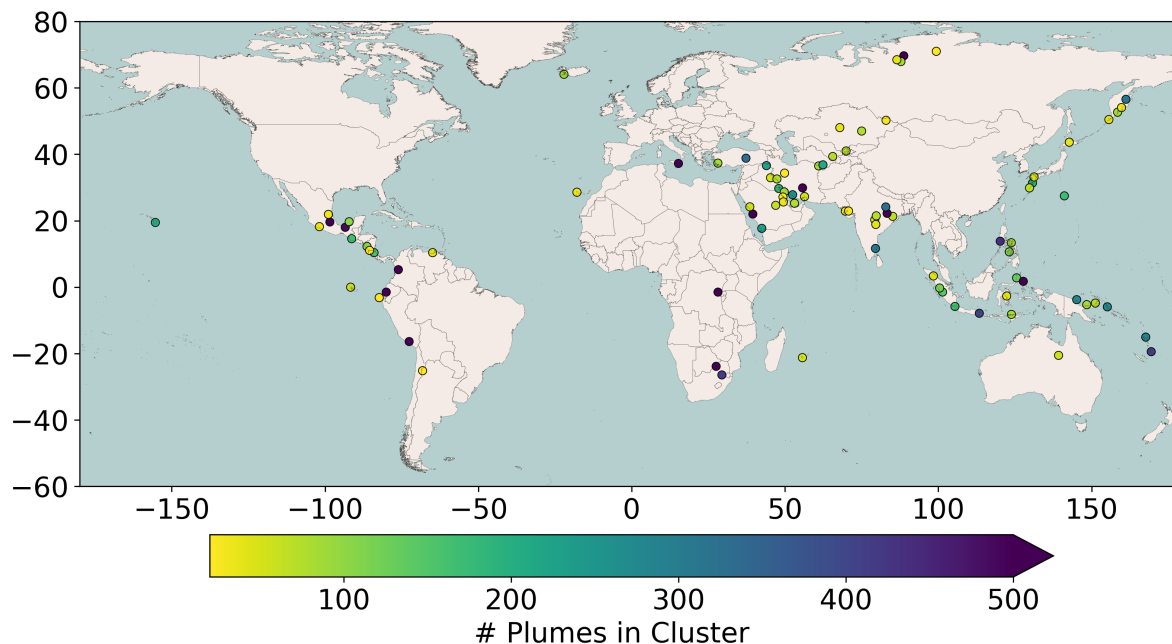


Figure 7. Locations of the centre of a cluster of plumes coloured by number of plumes in that cluster.

Using the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm (Ester et al., 1996), we grouped the detected plumes based on the coordinates of their maximum SO₂ observation (our proxy for origin). We set the clustering parameters to a maximum distance of 50 km between points and a minimum of 20 samples per cluster. This technique effectively identifies global areas with a high concentration of plumes. Figure 7 shows the centers of the 90 detected clusters, coloured by the number of plumes they contain, which further validates that our model successfully identifies the major SO₂ sources worldwide.

We observed regions on the global maps, most notably China (a location with numerous known SO₂ sources), where the model failed to detect expected plumes. Assuming the quality and volume of TROPOMI data are consistent globally, this deficiency likely stems from the plume detection model itself. We hypothesize that high background SO₂ concentrations in such regions may “hide” point sources, making plumes difficult to isolate. Since the model was trained on global data, it may struggle with these outlier scenarios. Future iterations should address this by specifically training on data that includes plumes from high-background regions like China.

Table 2 details the ten largest clusters of SO₂ plumes detected over the six-year study, alongside their probable sources. Volcanic activity dominates this list, accounting for seven out of the ten largest clusters. The remaining three clusters are associated with industrial activities: metal smelting, coal mining, and oil and gas operations. It is crucial to note that these clusters are often complex; some may contain multiple source types. For example, the Iztaccihuatl volcano cluster is located just south of Mexico City, meaning the cluster likely includes industrial SO₂ sources alongside the volcanic emissions.

Rank	Cluster Centre Coordinates	Nearest Likely Source	Source Type	# of Plumes	Median Emission Rate (kg hr ⁻¹)	Daily Persistence (%)
1	19.45°N, 98.54°W	Popocatépetl, Mexico	Volcano	1830	15,277	55.4
2	16.27°S, 72.17°W	Sabancaya, Peru	Volcano	1628	25,859	49.9
3	1.50°S, 80.01°W	Sangay, Ecuador	Volcano	1604	8582	42.7
4	5.26°N, 76.08°W	Nevado del Ruiz, Columbia	Volcano	1035	9716	37.2
5	69.65°N, 88.74°E	Norlisk, Russia	Metal Smelting	875	12,252	25.0
6	1.75°N, 127.71°E	Dukono, Indonesia	Volcano	782	9268	30.2
7	1.55°S, 28.17°E	Mount Nyiragongo/Nyamuragira, DRC	Volcano	668	7439	24.5
8	23.74°S, 27.43°E	Grootegeeluk Mine, S. Africa	Coal Mine & Power	612	12,769	26.0
9	37.24°N, 15.25°E	Mount Etna, Italy	Volcano	490	23,447	20.1
10	21.99°N, 39.51°E	Rabigh, Saudi Arabia	Oil and Gas	520	15,005	19.4

Table 2. Top ten largest clusters of plumes detected and their probable source, the total number of plumes over the six year study period, the median emission rate for all plumes in the cluster and the percentage of days a plume is detected in this region.

3.2 Coincidence with Volcanic Activity

Using the Smithsonian Institution's Global Volcanism Program (GVP) database of known eruptions since 1960 (GVP), we identified 227 active global eruptions during our study period. Comparing this list with our plume database revealed that 7,943 plumes were detected within a 50 km radius of an eruption during its active date. This confirms plume detection for 110 (49%) of these eruptions. The 117 missing detections are likely due to poor TROPOMI retrievals caused by high cloud cover or heavy aerosol loading. Figure 8 maps these GVP-reported eruptions, coloured by whether plumes were detected nearby (red) or not (green).

The plume detection algorithm struggles with very large plumes (> 1000 km) that often occur far downwind of major eruptions. Figure 9 illustrates this, showing a large SO₂ plume over the United Kingdom on August 24, 2024, originating from the Svartsengi volcanic system in Iceland (63.8°N, 22.38°W) six days earlier. The red outlines show the model splits this large plume into multiple smaller segments. This error likely results from the method used to break the TROPOMI swath into smaller images, which prevents the model from seeing the larger context of a plume that can span the entire swath.

3.3 TROPOMI SO₂ Activity Detection Flag

The TROPOMI Level 2 data assigns a flag to SO₂ pixels based on five categories: (0) no enhancement, (1) general SO₂ detection, (2) near a known volcano, (3) near a known anthropogenic source, and (4) a potential false positive due to a high solar zenith angle. Since a single plume often covers many pixels, we assign the final plume label only if over 80% of its pixels share the same TROPOMI flag (likely determined by spatial proximity to known sources); otherwise, the plume is labeled as a combined source. We found that the dominant category is a general SO₂ detection with no source attribution (43.5% of plumes), followed by volcanic plumes (30.8%) and known anthropogenic sources (14.7%). Figure 10 illustrates the geographical distribution of these categorized plumes.

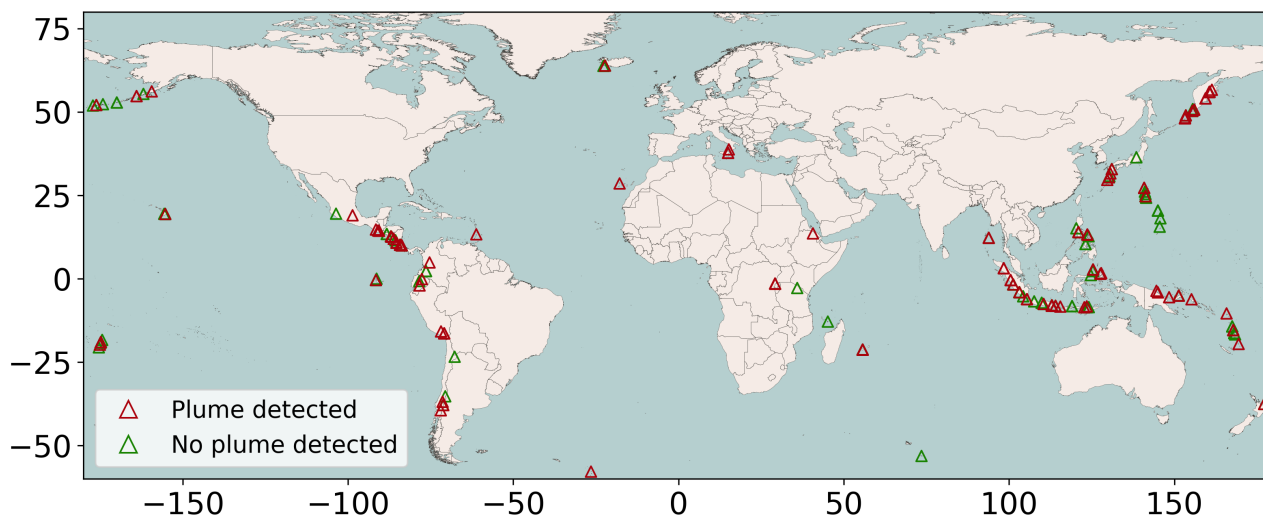


Figure 8. All eruptions reported between January 2019 and December 2024. Red triangles indicate plumes were detected within 50 km radius of the eruption during the eruptions dates. Green indicate no plume was detected.

The high number of unlabelled SO₂ detections is likely caused by plumes that have traveled downwind from their origin and are thus no longer close to the known volcanic or anthropogenic sources used for TROPOMI flagging. This explanation is strongly supported by the evidence of extensive volcanic outflow visible in both the geographical distribution of plumes (Figure 4) and the daily detection spikes (Figure 5).

The TROPOMI “No enhanced SO₂ detected” category accounts for 10.0% of all detections, but we believe many of these should be attributed to an actual source rather than being false positives. We demonstrate this by applying the clustering algorithm (as described in Section 3.1) to these “no enhanced detection” plumes. This allows us to extract meaningful groupings and calculate their proximity to clusters with known sources. Figure 11 displays all plumes in this category, highlighting 34 dense clusters (minimum 20 plumes within 150 km). Crucially, 27 of these 34 clusters overlap with clusters that have a known emission source, and the remaining seven are within 300 km. While these clusters account for 24% of the “no enhanced detection” plumes, many of the remaining plumes form a large, dispersed swath extending from 100°E to 180 °W around 10°S. This pattern aligns strongly with the downwind transport of SO₂ from the Taal Volcano eruption in the Philippines (120.99°E, 14.01°N) in April 2024 (detailed in Section 3.1).

This analysis demonstrates that while the source labelling provided in the TROPOMI files is informative, there is potential for our plume detection algorithm to complement the current flagging through masking and grouping plumes, facilitating analysis into repeated new detections. Although our plume detection algorithm does not attribute sources itself, it success-

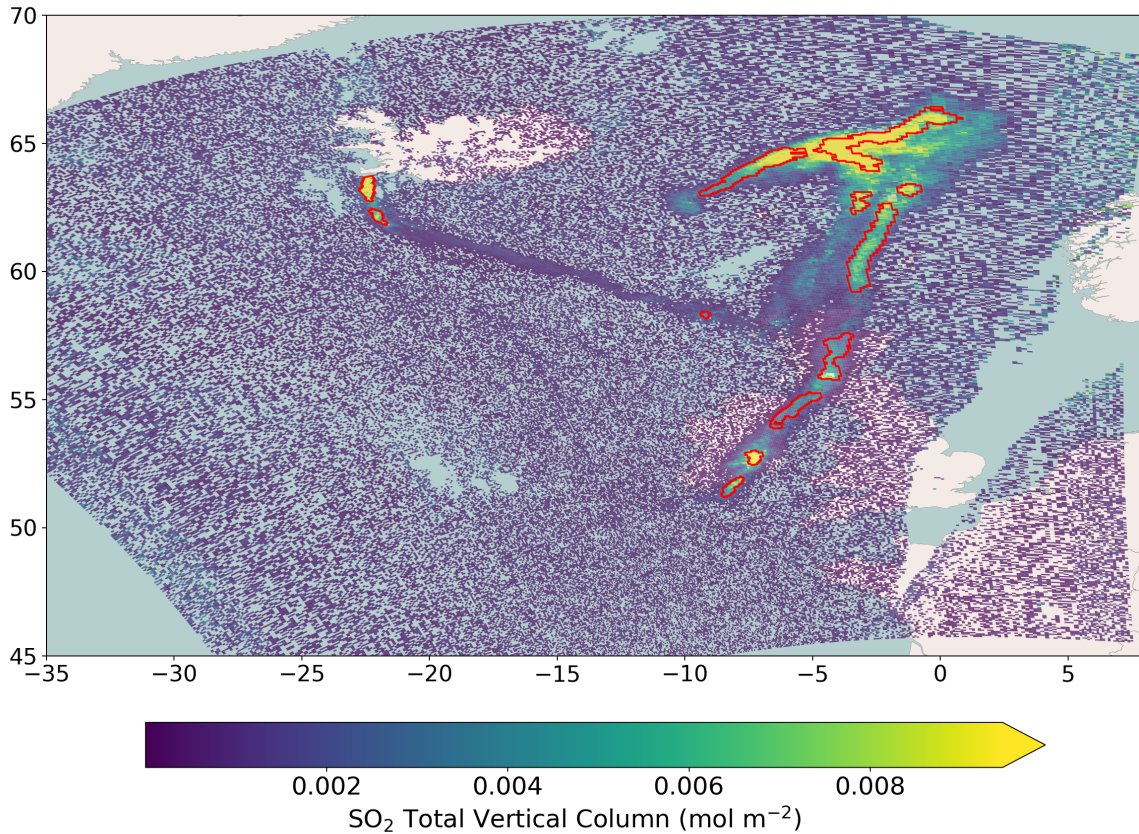


Figure 9. TROPOMI SO₂ total vertical column over the United Kingdom on 24th August 2024 showing a plume from the Svartsengi volcanic system in Iceland.

fully highlights geographical and temporal patterns that are crucial for identifying important, and potentially unlabelled, SO₂ emission locations.

4 Concluding Remarks

The segmentation model successfully processed approximately 31,000 TROPOMI Level 2 files from 2019 to 2024, demonstrating scalability and efficiency, with a rapid processing time of about 15 seconds per swath on a GPU. The methodology, which employed a 32×32 pixel rolling window and median probability overlap, effectively addressed issues related to plumes straddling image boundaries while ensuring the final plume shape was not constrained by the input image size. Although the raw performance metrics (Precision: 65.7%, Recall: 74%) may appear modest, they are negatively skewed by the high sensi-

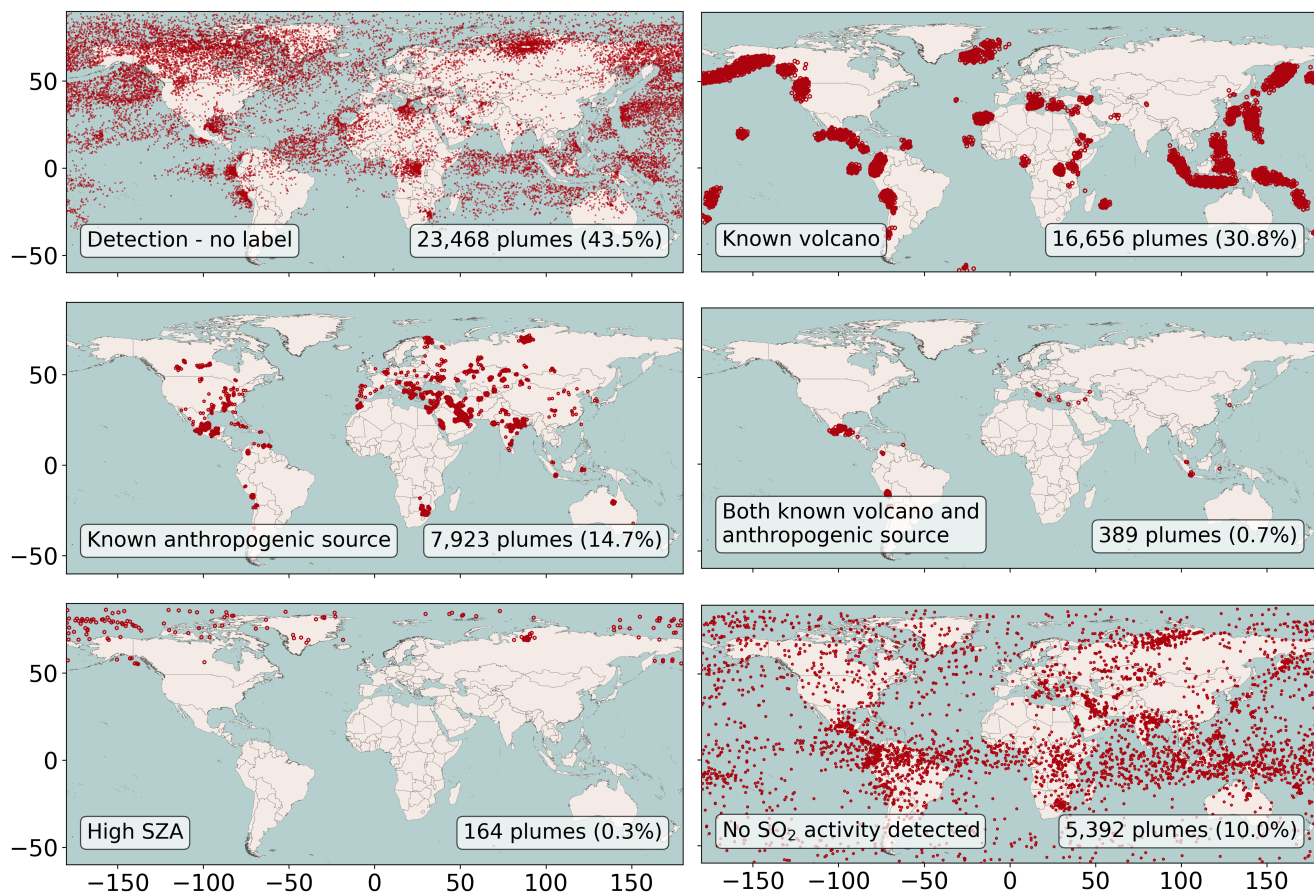


Figure 10. Global distribution of plume labels, 2019 - 2024, linked with the TROPOMI flag assigned to Level 2 SO₂ data.

tivity of using small 32×32 pixel images for validation. The model was deemed sufficient to establish a viable “first guess” emission database.

The global analysis confirmed that plumes cluster predictably around known volcanic and industrial hotspots. The detection of 53,993 plumes highlighted the critical role of episodic volcanic activity in SO₂ budgets, with significant annual spikes attributed to major events like the Raikoke and Ulawun eruptions in 2019. We also found that the TROPOMI Level 2 source flagging is incomplete; 43.5% of detections were unlabelled, a finding we attribute to plumes being transported far downwind from their designated source areas. Clustering analysis of the “No enhanced SO₂ detected” plumes (10.0% of detections) showed that many are, in fact, real sources, often corresponding to large, diluted outflow features like the plume from the 2024 Taal eruption.

However, two primary limitations must be addressed in future work. First, the algorithm struggles with very large plumes ($> \sim 1000$ km), often incorrectly segmenting them into multiple smaller plumes due to the inherent context limitation of the

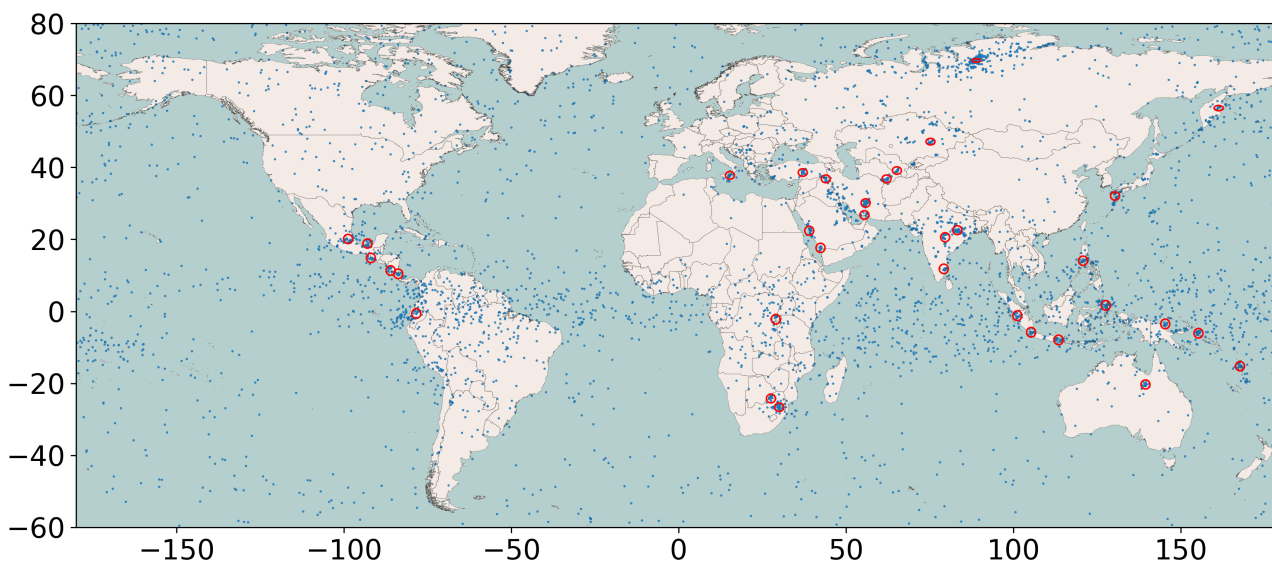


Figure 11. Location of all plumes labelled as “no enhanced detection” (blue), along with 34 clusters of more than 20 plumes within 150 km of each other (red) circles.

small image input method. Second, and more critically, there is a pronounced lack of plume detection over expected high-emission industrial regions like China and India. We hypothesise that the algorithm fails to isolate individual plumes against an excessively high SO_2 background concentration prevalent in these regions.

A crucial area for future advancement lies in integrating data from recently launched geostationary satellites. While the TROPOMI analysis relies on a single daily snapshot, instruments like the Geostationary Environment Monitoring Spectrometer (GEMS, covering Asia), the Tropospheric Emissions: Monitoring of Pollution (TEMPO, covering North America), and the future Sentinel-4 (covering Europe) offer sub-hourly or hourly measurements. This significantly higher temporal resolution would be transformative for this work, allowing for more robust plume tracking and improved wind field context for emission estimation. Furthermore, the hourly data would greatly enhance our ability to differentiate transient, high-flux plume events from persistent, high background SO_2 concentrations, thereby providing a pathway to potentially overcome the detection issues currently observed in heavily polluted industrial regions like China.

The application of machine learning is essential for achieving near real-time SO_2 source analysis, particularly for rapidly evolving natural hazards like volcanic eruptions. The immense processing speed of the segmentation model (sub-second per plume) is crucial for aviation safety, as SO_2 serves as a key proxy for hazardous volcanic ash. Machine learning automates the complex, time-consuming steps of plume boundary definition, geometric fitting (e.g., length), and emission rate calculation, transforming raw TROPOMI data into quantitative, actionable intelligence instantaneously. This rapid, objective assessment

has the potential to inform Volcanic Ash Advisory Centers and provide timely input for atmospheric transport models, greatly
295 enhancing warning systems and safety protocols.

Code availability. The plume detection code and plume annotation code can be requested from the authors.

Data availability. The SO₂ plume detection dataset is available from <https://zenodo.org/records/18302024>. TROPOMI SO₂ data are available from <https://dataspace.copernicus.eu/>

Appendix A: Training Dataset Creation

300 To create the training dataset for the U-Net model followed the following steps:

- We first used an image classification model developed by Finch et al. (2022) for NO₂ plume to create a dataset of images possibly containing an SO₂ plume. Although the model was not developed for SO₂, we relied on the similar characteristics of NO₂ and SO₂ (e.g. point source emission and relatively short lifetime) to have a first pass at reducing the number of images to manually check. We put all available SO₂ swaths from 2019 and 2020 through the classification
305 model for this first step.
- This first pass produced approximately 500 32 × 32 pixel images.
- We then developed a simple web application using Python and Plotly Dash (<https://dash.plotly.com>) which loaded each of these images one by one for manual verification and to draw a boundary around the plume. The web application also provided a RGB satellite image of the same geographical area as the TROPOMI image to help the user determine
310 if the plume was genuine or a likely false positive. A counter of the number of plumes verified was also provided. A screenshot of app is provided in figure A1.
- To draw the boundary around the plume, the user uses the mouse cursor to draw a line around what they believe is the plume boundary. This line is then saved and converted to both latitude and longitude coordinates and TROPOMI swath index coordinates and a plume mask is produced relating to the image.
- 315 – This process was repeated after numerous iterations of training and checking the plume detection model to refine the output.

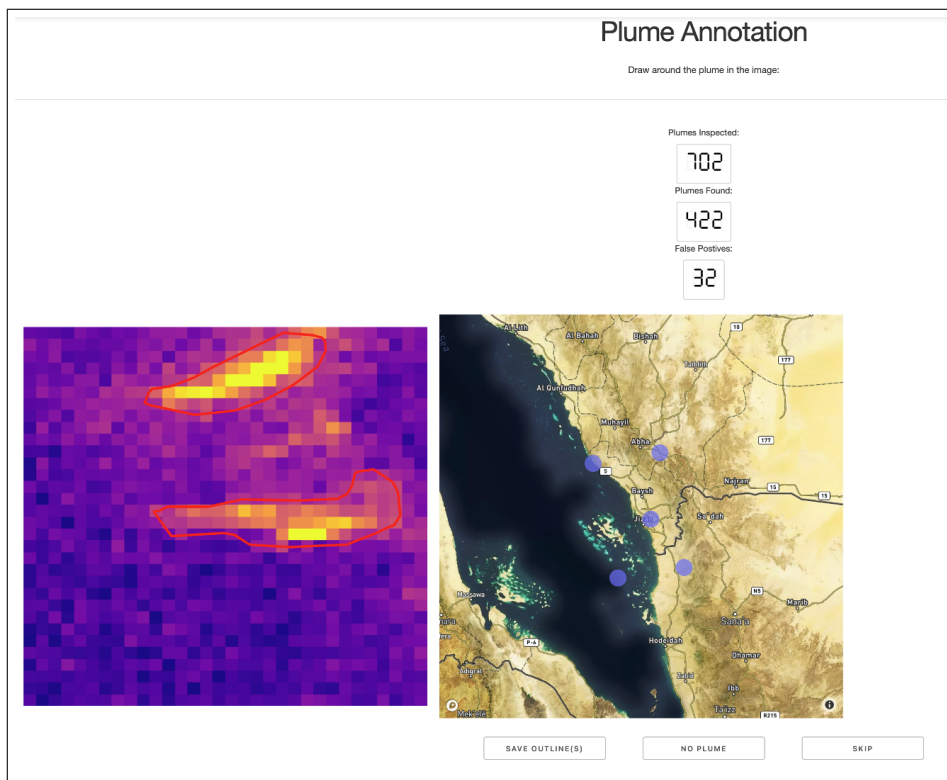


Figure A1. A screenshot of the plume annotation web app developed to draw boundaries around plumes of SO₂ in TROPOMI data. The image and the left shows the TROPOMI data and the image on the right shows a RGB image of the same ground location to use as a sanity check for the plume.

Appendix B: Major Volcanic Outflow

Figure B1 shows the plumes detected near likely volcanic eruptions, coloured by the days elapsed since the first eruption was detected. These show how these plumes are likely outflow from these large eruptions which are likely being detected on
 320 subsequent days as they are transported downwind.

Appendix C: Plume Dataset Description

The resulting plume dataset described in this paper is available for use and contains the following information:

- Date and time of the TROPOMI swath
 - Latitude and longitude of the maximum SO₂ value within the plume (degrees north and west)
 - X and Y Index on the TROPOMI swath of location of the maximum SO₂ value within the plume
- 325

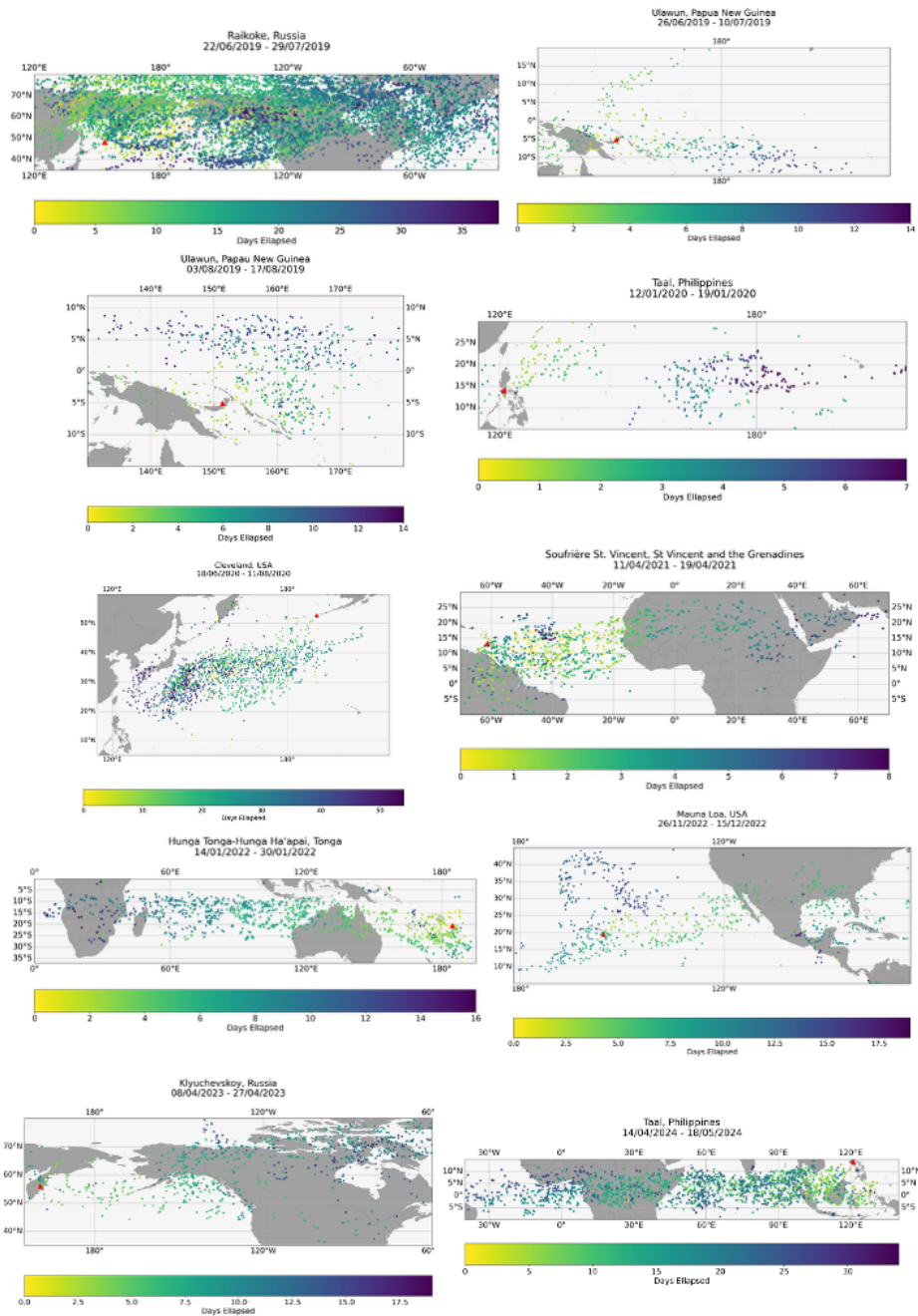


Figure B1. Plumes likely associated with major volcanic eruptions through the study period. The colouring represents days elapsed since first eruption detected.

- Value of the maximum SO₂ concentration within the plume (mol m⁻²)
- Indices on the TROPOMI swath of the bounding box surrounding the plume
- Plume mass enhancement from the background (kg)
- Latitude and longitudes of the plume border (degree north and west)
- 330 – Indices on the TROPOMI swath of the plume border
- Data on the fitted ellipse (x and y of ellipse centre, width, height and angle)
- Median, minimum, maximum, standard deviation and standard error of wind speed (ms⁻¹)
- Wind direction (degrees)
- Plume length (m)
- 335 – Plume emission estimate (kg hr⁻¹)
- Array of pixel areas within the bounding box (m²)
- U and V wind fields within the bounding box
- Name of the original TROPOMI file used

340 *Author contributions.* DPF and PIP designed the research; DPF prepared the calculations; DPF and PIP analysed the results and wrote the paper.

Competing interests. There are no competing interests present.

Acknowledgements. Douglas P. Finch and Paul I. Palmer gratefully acknowledge funding from the ESA for the EOPlumes project.

References

- Arellano, S., Galle, B., Apaza, F., Avard, G., Barrington, C., Bobrowski, N., Bucarey, C., Burbano, V., Burton, M., Chacón, Z., Chigna, G.,
345 Clarito, C. J., Conde, V., Costa, F., De Moor, M., Delgado-Granados, H., Di Muro, A., Fernandez, D., Garzón, G., Gunawan, H., Haerani,
N., Hansteen, T. H., Hidalgo, S., Inguaggiato, S., Johansson, M., Kern, C., Kihlman, M., Kowalski, P., Masias, P., Montalvo, F., Möller,
J., Platt, U., Rivera, C., Saballos, A., Salerno, G., Taisne, B., Váscónez, F., Velásquez, G., Vita, F., and Yalire, M.: Synoptic analysis of a
decade of daily measurements of SO₂ emission in the troposphere from volcanoes of the global ground-based Network for Observation of
Volcanic and Atmospheric Change, *Earth System Science Data*, 13, 1167–1188, <https://doi.org/10.5194/essd-13-1167-2021>, 2021.
- 350 Bazalgette Courrèges-Lacoste, G., Sallusti, M., Bulsa, G., Bagnasco, G., Veihelmann, B., Riedl, S., Smith, D. J., and Maurer, R.: The Copernicus Sentinel 4 mission: a geostationary imaging UVN spectrometer for air quality monitoring, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 10423, 1042307, <https://doi.org/10.1117/12.2282158>, ISBN: 9781510613102, 2017.
- Bokstaller, J., She, Y., Fu, Z., and Macrì, T.: Segmentation of Roads in Satellite Images using specially modified U-Net CNNs, <https://doi.org/10.48550/arXiv.2109.14671>, arXiv:2109.14671 [cs], 2021.
- 355 Bradski, G.: *The OpenCV Library*, Dr. Dobb's Journal of Software Tools, 2000.
- Brenot, H., Theys, N., Clarisse, L., van Geffen, J., van Gent, J., Van Roozendaal, M., van der A, R., Hurtmans, D., Coheur, P.-F., Clerbaux, C., Valks, P., Hedelt, P., Prata, F., Rasson, O., Sievers, K., and Zehner, C.: Support to Aviation Control Service (SACS): an online service for near-real-time satellite monitoring of volcanic plumes, *Natural Hazards and Earth System Sciences*, 14, 1099–1123, <https://doi.org/10.5194/nhess-14-1099-2014>, 2014.
- 360 Carn, S. A., Fioletov, V. E., McLinden, C. A., Li, C., and Krotkov, N. A.: A decade of global volcanic SO₂ emissions measured from space, *Scientific Reports*, 7, 44095, <https://doi.org/10.1038/srep44095>, number: 1, 2017.
- Chance, K., Liu, X., Miller, C. C., Abad, G. G., Huang, G., Nowlan, C., Souri, A., Suleiman, R., Sun, K., Wang, H., Zhu, L., Zoogman, P., Al-Saadi, J., Antuña-Marrero, J.-C., Carr, J., Chatfield, R., Chin, M., Cohen, R., Edwards, D., Fishman, J., Flittner, D., Geddes, J., Grutter, M., Herman, J. R., Jacob, D. J., Janz, S., Joiner, J., Kim, J., Krotkov, N. A., Lefer, B., Martin, R. V., Mayol-Bracero, O. L., Naeger, A.,
365 Newchurch, M., Pfister, G. G., Pickering, K., Pierce, R. B., Cárdenas, C. R., Saiz-Lopez, A., Simpson, W., Spinei, E., Spurr, R. J. D., Szykman, J. J., Torres, O., and Wang, J.: TEMPO Green Paper: Chemistry, physics, and meteorology experiments with the Tropospheric Emissions: monitoring of pollution instrument, in: *Sensors, Systems, and Next-Generation Satellites XXIII*, vol. 11151, pp. 56–67, SPIE, <https://doi.org/10.1117/12.2534883>, 2019.
- Driscoll, C. T., Lawrence, G. B., Bulger, A. J., Butler, T. J., Cronan, C. S., Eagar, C., Lambert, K. F., Likens, G. E., Stoddard, J. L., and
370 Weathers, K. C.: Acidic Deposition in the Northeastern United States: Sources and Inputs, Ecosystem Effects, and Management Strategies, *BioScience*, 51, 180, [https://doi.org/10.1641/0006-3568\(2001\)051\[0180:ADITNU\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0180:ADITNU]2.0.CO;2), 2001.
- EDGARv8.1: EDGAR - The Emissions Database for Global Atmospheric Research, https://edgar.jrc.ec.europa.eu/emissions_data_and_maps.
- Ester, M., Kriegel, H.-P., Sander, J., and Xu, X.: A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, *Second International Conference on Knowledge Discovery and Data Mining (KDD'96)*, pp. 226–331, 1996.
- 375 Filatov, D. and Yar, G. N. A. H.: Forest and Water Bodies Segmentation Through Satellite Images Using U-Net, <https://doi.org/10.48550/arXiv.2207.11222>, arXiv:2207.11222 [cs], 2022.
- Finch, D. P., Palmer, P. I., and Zhang, T.: Automated detection of atmospheric NO₂ plumes from satellite data: a tool to help infer anthropogenic combustion emissions, *Atmospheric Measurement Techniques*, 15, 721–733, <https://doi.org/10.5194/amt-15-721-2022>, 2022.

- 380 Galle, B., Johansson, M., Rivera, C., Zhang, Y., Kihlman, M., Kern, C., Lehmann, T., Platt, U., Arellano, S., and Hidalgo, S.: Network for Observation of Volcanic and Atmospheric Change (NOVAC)—A global network for volcanic gas monitoring: Network layout and instrument description, *Journal of Geophysical Research: Atmospheres*, 115, 2009JD011 823, <https://doi.org/10.1029/2009JD011823>, 2010.
GVP: Global Volcanism Program, <https://volcano.si.edu/>.
- 385 Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J.-i., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), *Geoscientific Model Development*, 11, 369–408, <https://doi.org/10.5194/gmd-11-369-2018>, 2018.
- Khalaf, E. M., Mohammadi, M. J., Sulistiyani, S., Ramírez-Coronel, A. A., Kiani, F., Jalil, A. T., Almulla, A. F., Asban, P., Farhadi, M.,
390 and Derikondi, M.: Effects of sulfur dioxide inhalation on human health: a review, *Reviews on Environmental Health*, 39, 331–337, <https://doi.org/10.1515/reveh-2022-0237>, 2024.
- Kim, J., Jeong, U., Ahn, M.-H., Kim, J. H., Park, R. J., Lee, H., Song, C. H., Choi, Y.-S., Lee, K.-H., Yoo, J.-M., Jeong, M.-J., Park, S. K., Lee, K.-M., Song, C.-K., Kim, S.-W., Kim, Y. J., Kim, S.-W., Kim, M., Go, S., Liu, X., Chance, K., Miller, C. C., Al-Saadi, J., Veihelmann, B., Bhartia, P. K., Torres, O., Abad, G. G., Haffner, D. P., Ko, D. H., Lee, S. H., Woo, J.-H., Chong, H., Park, S. S., Nicks, D., Choi, W. J.,
395 Moon, K.-J., Cho, A., Yoon, J., Kim, S.-k., Hong, H., Lee, K., Lee, H., Lee, S., Choi, M., Veeffkind, P., Levelt, P. F., Edwards, D. P., Kang, M., Eo, M., Bak, J., Baek, K., Kwon, H.-A., Yang, J., Park, J., Han, K. M., Kim, B.-R., Shin, H.-W., Choi, H., Lee, E., Chong, J., Cha, Y., Koo, J.-H., Irie, H., Hayashida, S., Kasai, Y., Kanaya, Y., Liu, C., Lin, J., Crawford, J. H., Carmichael, G. R., Newchurch, M. J., Lefer, B. L., Herman, J. R., Swap, R. J., Lau, A. K. H., Kurosu, T. P., Jaross, G., Ahlers, B., Dobber, M., McElroy, C. T., and Choi, Y.:
400 New Era of Air Quality Monitoring from Space: Geostationary Environment Monitoring Spectrometer (GEMS), *Bulletin of the American Meteorological Society*, 101, E1–E22, <https://doi.org/10.1175/BAMS-D-18-0013.1>, 2020.
- Mukhopadhyay, A., Oksuz, I., Bevilacqua, M., Dharmakumar, R., and Tsiftaris, S. A.: Unsupervised Myocardial Segmentation for Cardiac MRI, in: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, edited by Navab, N., Hornegger, J., Wells, W. M., and Frangi, A. F., vol. 9351, pp. 12–20, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-319-24574-4_2, series Title: Lecture Notes in Computer Science, 2015.
- 405 Pan, Z., Xu, J., Guo, Y., Hu, Y., and Wang, G.: Deep Learning Segmentation and Classification for Urban Village Using a Worldview Satellite Image Based on U-Net, *Remote Sensing*, 12, 1574, <https://doi.org/10.3390/rs12101574>, number: 10, 2020.
- Romahn, F., Pedergnana, M., Loyola, D., Apituley, A., Sneep, M., and Veeffkind, J.: Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Sulphur Dioxide SO₂, Tech. Rep. 2.05, DLR.
- Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, in: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, edited by Navab, N., Hornegger, J., Wells, W. M., and Frangi, A. F., pp. 234–241, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-319-24574-4_28, 2015.
- 410 Smith, R. A.: *Air and rain : the beginnings of a chemical climatology*, London : Longmans, Green, and Co., 1872.
- Soulie, A., Granier, C., Darras, S., Zilbermann, N., Doumbia, T., Guevara, M., Jalkanen, J.-P., Keita, S., Liousse, C., Crippa, M., Guizzardi, D., Hoesly, R., and Smith, S.: Global Anthropogenic Emissions (CAM5-GLOB-ANT) for the Copernicus Atmosphere Monitoring Service
415 Simulations of Air Quality Forecasts and Reanalyses, preprint, ESSD – Atmosphere/Energy and Emissions, <https://doi.org/10.5194/essd-2023-306>, 2023.

- Srivastava, R. K., Jozewicz, W., and Singer, C.: SO₂ scrubbing technologies: A review, *Environmental Progress*, 20, 219–228, <https://doi.org/10.1002/ep.670200410>, eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/ep.670200410>, 2001.
- 420 Theys, N., De Smedt, I., Yu, H., Danckaert, T., van Gent, J., Hörmann, C., Wagner, T., Hedelt, P., Bauer, H., Romahn, F., Pedernana, M., Loyola, D., and Van Roozendael, M.: Sulfur dioxide retrievals from TROPOMI onboard Sentinel-5 Precursor: algorithm theoretical basis, *Atmospheric Measurement Techniques*, 10, 119–153, <https://doi.org/10.5194/amt-10-119-2017>, 2017.
- Theys, N., Hedelt, P., De Smedt, I., Lerot, C., Yu, H., and Van Roozendael, M.: S5P/TROPOMI SO₂ ATBD, Tech. Rep. 2.5.0, BIRA, 2023.
- Ulmas, P. and Liiv, I.: Segmentation of Satellite Imagery using U-Net Models for Land Cover Classification, <https://doi.org/10.48550/arXiv.2003.02899>, arXiv:2003.02899 [cs], 2020.
- 425 Veefkind, J., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H., de Haan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, *Remote Sensing of Environment*, 120, 70–83, <https://doi.org/10.1016/j.rse.2011.09.027>, 2012.
- Vernier, J.-P., Aubry, T. J., Timmreck, C., Schmidt, A., Clarisse, L., Prata, F., Theys, N., Prata, A. T., Mann, G., Choi, H., Carn, S., Rigby, R.,
430 Loughlin, S. C., and Stevenson, J. A.: The 2019 Raikoke eruption as a testbed used by the Volcano Response group for rapid assessment of volcanic atmospheric impacts, *Atmospheric Chemistry and Physics*, 24, 5765–5782, <https://doi.org/10.5194/acp-24-5765-2024>, 2024.
- Watson-Parris, D., Christensen, M. W., Laurenson, A., Clewley, D., Gryspeerd, E., and Stier, P.: Shipping regulations lead to large reduction in cloud perturbations, *Proceedings of the National Academy of Sciences*, 119, e2206885 119, <https://doi.org/10.1073/pnas.2206885119>, 2022.
- 435 Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A.: Learning Deep Features for Discriminative Localization, <https://doi.org/10.48550/arXiv.1512.04150>, arXiv:1512.04150 [cs], 2015.