Response to Review
Anonymous Reviewer #3

Finch et al. presented an SO2 data set based on machine learning identification of plumes and emission estimation from TROPOMI. It is an interesting study, and I believe the work could be published after revision. The paper is well written and structured. Overall, the figures are of sufficient quality.

My main reservation is that the SO2 fluxes derived are not evaluated against other available estimates. Over the last years, several studies have reported SO2 top-down estimates using TROPOMI, but the main author does not cite those papers. In particular, Fioletov et al. 2023 (https://doi.org/10.5194/essd-15-75-2023) provides SO2 emission estimates and I would like to see a comparison between the emissions from this work and the results from Fioletov. This can be done for several representative SO2 sources (anthropogenic and volcanic), with stable emissions. This should come with a discussion of pro and cons for the presented method. Apart from that, I agree with all the comments raised by Pascal Hedelt (some are repeated below) and the author should address them in the replies and revised manuscript.

We thank the reviewer for their comments and have addressed their suggestions below.

Introduction

Line 33: "... with recent years showing lower values". Please add a reference.

We have moved the reference to Soulie et al (2023) to the end of the sentence to include this statement.

Line 57: NOVAC. Please add a reference to Galle et al. 2010 https://doi.org/10.1029/2009JD011823

This has been added.

Methodology: section 2.1

-Which SO2 product version was used? Recently, the TROPOMI SO2 product has switched to the COBRA algorithm which is more sensitive to weak SO2 emissions. Did the author use this data? If not, why not?

We used version 3 of the offline TROPOMI data, using the DOAS retrieval for the SO_2 product. We did not use the COBRA retrieval for this paper because, at the time of data download and processing, the full COBRA dataset was not yet available. Although the complete dataset has since become been released and published evaluaions indicate that it performs well, incorporating it here would require re-processing the entire analysis at a computational cost beyond the scope of this paper.

-In line with P. Hedelt comment, more information must be given on which SO2 column product was used (the main VCD and/or the 1,7,15km VCD product) and what is the impact of this choice on the final result.

This is addressed in the response two points below this one.

-A reference to the main papers, ATBD and product read me file should be added.

These have now been added to the manuscript.

-The quality flag >0.5 applies to the main VCD product which assumes an SO2 profile from pollution. This flag removes much of the cloudy pixels which are still very useful for volcanic events where the SO2 plume lies over clouds (in this case, the 1,7,15km VCD product are more appropriate). This is not assessed or discussed in the paper.

We have added the following statement:

Because this study uses the full vertical column density (VCD) rather than the 1, 7 or 15km layer products, we acknowledge that some potential SO_2 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.

Commented [PP1]: unless this acronym was introduced earlier.

Methodology: section 2.2

-line 100: in line with P. Hedelt, the 'manual creation of a precise plume mask' deserves a thorough description.

We have now added a new appendix that includes a detailed description of how we created the training dataset of plumes and plume masks.

-It is not clear to me what is the added-value of the proposed plume detection compared to the detection flag. The selective detection of SO2 from a hyperspectral instrument like TROPOMI is relatively straightforward, and because the SO2 background is negligible it is easy to identify the plumes. The proposed method would perform better for species like CH4 for which the background level is significant.

This method defines the outlines of plumes whereas the detection flag is on a per pixel basis. Further analysis is needed with the detection flag to determine boundary thresholds or distinct plumes. We acknowledge that the detection flag does a good job a lot of the time and is adequate for many uses but do not consider this a reason not to try and develop other methods of plume detection.

The ESA EOPlumes project, from which this study originates, included parallel analyses for NO_2 and CH_4. We chose to publish the SO_2 component first because it was the most mature and provided the clearest demonstration of how machine learning methods can accelerate the processing of EO data.

Methodology: section 2.3

-Line 35: typo: "To estimate the the emission"

This has been corrected.

-it would be informative to compare the elipse main axis direction with the wind direction used to estimate the SO2 emissions. Do they compare well?

We did try this and the results were mixed. As with the other issues with the emission calculations, the wind is likely not representative if the plume is at a high altitude. At the request of another reviewer, we have reduced the section on emissions down to a small example of what could potentially be done so the issues with the wind are no longer relevant to this study.

Section 3

-line 183: typo – a question mark appears in (Vernier et al., 2024;?).

Thank you for spotting this formatting error. This now reads "(Vernier et al., 2024, and references therein)"

-l184: about Peak I, if it is related to Norilsk, I don't see why it appears only for this year.

We agree it is unusual behaviour and we have been unable to find a clear reason for this, but we consider it interesting enough to include in this study. Any data on change in production in the area, which is the most likely cause, is unavailable to us.

It has also been pointed out that the spike in detections also corresponds to an eruption in eastern Russia which has now been included in the manuscript.

-l195: about the high background SO2 concentrations. Over China, the SO2 levels are quite low. Indeed, there have been many regulations on SO2 emissions in China, and I think this statement is not true. Later, the main author argues on the high SO2 levels in China based on EDGAR inventory but it is not clear to me if EDGAR is up-to-date regarding the SO2 emissions level over China or not.

While there have been many successful regulations implemented in China, and SO_2 levels have reduced dramatically in the past two decades, the levels of SO_2 are still higher than compared with many other countries. We agree the EDGAR inventory may not be up to date, we were unable to find any other suitable inventory that covers the study period. We do state in the paper that this is our hypothesis and not drawn from thorough analysis.

-section 3.5: the presentation of the emission database is minimal. It consists mainly of Figure 12 which is not very informative. I would like to see at least maps of emissions (global, regional, per emission type, etc). The comparison with EDGAR is also weak in my opinion. The author mainly describes why it is presumably not possible to compare. As a reader, it is not clear to me what this section is about.

At the request of another reviewer, we have reduced this section down to an example of an emission calculation to show what could potentially be done as a next step in this project.

Conclusions

The last sentence about the usefulness of the approach for the VAACs is doubtful. A simple VCD threshold mask (or detection flag) is enough to isolate the plumes.

We acknowledge that the final sentence uses some hyperbole and have rewritten it emphasise the potential usefulness of this product.

A threshold mask would be on a per pixel basis and therefore need further analysis to determine clusters of pixels and what constitutes a single plume or background noise. This method has been designed to try and eliminate the need for that step on every detection and provide a quick and simple assessment of a plume.

Data availability

The dataset should be available as supplement or in a data repository

We have uploaded the data to the following repository**:** https://zenodo.org/records/18302024 . This has been added to the manuscript. These data are open access.

Acknowledgements

Please provide information on the ESA project supporting this work.

We have added the name of the project (EOPlumes) into the acknowledgements.