S5P/TROPOMI COBRA SO_2 ALGORITHM SO2CBR

VALIDATION REPORT

document number: S5P- BIRA-L2-VR-SO2CBR

issue : 2.0.0
date : 2023 12 20

Document approval record

	digital signature
prepared:	Nicolas Theys
Checked:	
approved PI:	
approved PM:	

Document change record

issue	date	item	comments
1.0.0	2022-09-14	All	First complete document version
2.0.0	2023-12-20	All	Update of the document for SO2CBR version 2

Contents

Document approval record 2
Document change record 3
List of figures 5
List of tables 6
1 Introduction 7
1.1 Identification 7
1.2 Purpose and objective 7
1.3 Document overview 7
1.4 Acknowledgements 7
2 Applicable and reference documents 8
2.1 Applicable documents 8
2.2 Standard documents 8
2.3 Reference documents 8
2.4 Electronic references 8
3 Terms, definitions and abbreviated terms 9
3.1 Terms and definitions 9
3.2 Acronyms and abbreviations 9
4 Introduction 10
4.1 The SO2CBR L2 product 10
4.2 Data quality requirements and validation 10
5 Slant column verification 12
6 Validation of SO2CBR vertical column density product 14
6.1 Comparison with TROPOMI SO_2 products 14
6.2 Comparison with ground-based data. 15
6.2.1 TROPOMI COBRA versus MAX-DOAS measurements 15
6.2.2 TROPOMI COBRA versus Pandora measurements 17
7 Improvement of the SO2CBR v2 product quality 19
8 Conclusions 21

List of figures

Figure 1: Comparison of SO_2 slant column data for one week (1-7 June 2021) for the prototype algorithm-BIRA (top) and the PAL implemented algorithm (center). The SO_2 VCDs difference map of PAL minus BIRA is show on the lowest panel. 13
Figure 2: Comparison of seasonal mean SO_2 columns for September to November 2019 retrieved from TROPOMI DOAS, COBRA and PCA algorithms (from top to bottom). Consistent pixel selection criteria, gridding and retrieval settings are applied. For all four data sets, a fixed AMF of 0.4 is applied. 15
Figure 3: (left) Comparison of monthly mean SO_2 columns from MAX-DOAS and TROPOMI COBRA for (a) Xianghe, and (b) Mohali. The grey and pale red dots correspond to the individual days. (right) Scatter plots of monthly mean SO_2 columns of TROPOMI COBRA vs MAX-DOAS observations. Error bars are the standard errors on the monthly average SO_2 columns. The correlation coefficient and slope of the regression line are given as an inset for each plot. 16
Figure 4: The ground measurements stations for the Pandonia Global Network. Image from PGN website [URL03], last access 30-06-2022. 17
Figure 5: Comparison of TROPOMI COBRA SO2 VCDs to preliminary Pandora measurements (daily means) at Mexico City and Wakkerstroom from September 2019 to August 2021. The following selection criteria were applied: distance < 25km, CF<0.2, time window +/- 1h around overpass. 18
Figure 6 Comparison of TROPOMI COBRA SO2 VCDs from (top) version 1 and (bottom) version 2, averaged over the period from 9 to 16 April 2021. The large SO2 plume observed in the tropical band is from the eruption of La Soufrière. 19
Figure 7 Comparison of TROPOMI SO2 VCD monthly averages (left: April 2021, right October 2021) from COBRA version 1 (top) and COBRA version 2 (bottom). The extended SO2 plume in October 2021 near the West coast of Africa is due to the eruption of Cumbre Vieja. 20

List of tables

Table 1: Data quality target for the Sentinel-5 Precursor TROPOMI L2 SO2 product.................................**10**

1 Introduction

1.1 Identification

This document describes the verification and validation of SO2CBR L2 data by means of comparisons with independent datasets. The SO2CBR version described in this document is version 02.00.00.

1.2 Purpose and objective

The purpose of the document to assess the quality of SO2CBR derived SO2 columns from S5P measurements. The objective is to give insight in the current quality of the results and remaining work to be done.

1.3 Document overview

Chapter 2 lists applicable and reference documents. Chapter 3 gives relevant terms and definitions. Chapter 4 introduces the SO2CBR product, data quality requirements and validation challenges. The actual data comparison is presented in chapters 5,6 & 7. Compared to the processor version 01.00.00, the algorithm changes in v02.00.00 are mostly for specific conditions, namely largely volcanic eruptions, and observations at high latitudes. These are covered by Chapter 7. Results in Chapters 5 and 6 are based on SO2CBR version 01.00.00 (but the findings are valid for version 02.00.00 as well). Conclusions are given in Chapter 8.

1.4 Acknowledgements

The authors would like to thank the following people for their kind assistance in the presented work: Alexander Cede and Martin Tiefengraber from Luftblick for making available ground-based SO_2 observations from the Pandonia Global Network.

2 Applicable and reference documents

2.1 Applicable documents

There are no applicable documents.

2.2 Standard documents

There are no standard documents.

2.3 Reference documents

[RD01] S5p L2 COBRA paper: Theys, N., Fioletov, V., Li, C., De Smedt, I., Lerot, C., McLinden, C., Krotkov, N., Griffin, D., Clarisse, L., Hedelt, P., Loyola, D., Wagner, T., Kumar, V., Innes, A., Ribas, R., Hendrick, F., Vlietinck, J., Brenot, H., and Van Roozendael, M.: A sulfur dioxide Covariance-Based Retrieval Algorithm (COBRA): application to TROPOMI reveals new emission sources, Atmos. Chem. Phys., 21, 16727–16744, https://doi.org/10.5194/acp-21-16727-2021, 2021.

[RD02] Sentinel-5 precursor/TROPOMI Level 2 Algorithm Theoretical Basis Document Sulphur Dioxide SO2, **source:** BIRA-IASB; **ref:** S5P-BIRA-L2-400E-ATBD; **url**: https://sentinel.esa.int/documents/247904/2476257/Sentinel-5P-ATBD-SO2-TROPOMI

[RD03] Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products; **ref:** S5P-MPC-IASB-ROCVR; **issue:** 21.01.01; **date:** 2023-12-19.

2.4 Electronic references

[URL01] https://www.s5p-pal.com/

[URL02] https://uv-vis.aeronomie.be/geoserver/cobra/

[URL03] https://www.pandonia-global-network.org/

[URL04] http://mpc-vdaf.tropomi.eu/

3 Terms, definitions, and abbreviated terms

Terms, definitions, and abbreviated terms that are used in the development program for the TROPOMI L2 data processors are described [RD02]. Terms, definitions, and abbreviated terms that are specific for this document can be found below.

3.1 Terms and definitions

No additional terms at this point.

3.2 Acronyms and abbreviations

AMF	Air Mass Factor
BIRA	Royal Belgian Institute for Space Aeronomy
CF	Cloud Fraction
COBRA	Covariance-Based Retrieval Algorithm
DOAS	Differential Optical Absorption Spectroscopy
DU	Dobson Unit
ESA	European Space Agency
L2	Level-2
MAX-DOAS	Multi-axis DOAS
PAL	Product Algorithm Laboratory
PCA	Principal Component Analysis
PGN	Pandonia Global Network
S5P	Sentinel-5 Precursor
SCD	Slant Column Density
SZA	Solar Zenith Angle
SO2	Sulfur dioxide
TROPOMI	Tropospheric Monitoring Instrument
VCD	Vertical Column Density

4 Introduction

4.1 The SO2CBR L2 product

The SO2CBR algorithm for the derivation of L2 vertical SO_2 column data was developed in the framework of the ESA-funded S5P-PAL project [URL01], coordinated by S&T. The algorithm follows a so-called Covariance-Based Retrieval Algorithm (COBRA) scheme. A full description of the algorithm details is given in [RD01]. The current version of SO2CBR is closely linked to the operational SO2 algorithm [RD02] and shares several algorithm modules. Compared to the operational SO_2 algorithm based on Differential Optical Absorption Spectroscopy (DOAS), the COBRA fits an SO_2 slant column without the need for a post-processing offset correction. The retrieval is based on a measurement error covariance matrix to fully represent the SO_2-free radiance variability, so that the SO_2 slant column density is the only retrieved parameter of the fitting algorithm. The strength of the method is that the covariance matrix is adjusted dynamically and separately for each TROPOMI orbit, each row and scanline segment. This complexity warrants the results to be significantly improved essentially for low SO_2 columns in two aspects. First, zonal and local biases (on the order of 0.2 DU in the operational VCD product) are significantly reduced allowing SO_2 emission source areas to be studied in much greater detail (see maps in [URL02]). Second, the noise on the data is reduced by a factor of 2.

We note that the SO2CBR algorithm applies only to the slant columns and only in fitting window 1. For large SO_2 columns, the windows 325-335 nm and 360-390 nm (windows 2 and 3) are used, and are directly obtained from the operational product. To convert the slant columns into vertical columns, air mass factor calculations are needed, and a slightly adapted version of the look-up-table approach used in the operation algorithm [RD02] has been implemented.

4.2 Data quality requirements and validation

From previous validation exercises, targets for key quality indicators of S5P SO2 L2 data product have evolved and have been formulated in [RD03]. For SO_2 column data, the targets are as reported in Table 1.

Quantity	Requirement: uncertainty due to Systematic effects	Requirement: uncertainty due to Random effects
Total SO_2 column	0.5 DU	1 DU
Enhanced SO_2 column (SCD<1.5 DU)	0.5 DU	1 DU
Enhanced SO_2 column (SCD>1.5 DU)	30%	30%

Table 1: Data quality target for the Sentinel-5 Precursor TROPOMI L2 SO2 product.

Note that a distinction is made between volcanic SO_2 conditions (referred as 'Enhanced' in Table 1) and SO_2 pollution scenarios in the boundary layer ('Total'). For both operational and COBRA products, the requirement on the bias is generally fulfilled for volcanic plumes, while the dispersion can exceed slightly the requirement on the random component of the uncertainty, which is not considered as a substantial restriction of the data quality. For polluted scenes, the operational SO_2 product fulfill the requirements overall but these are rather lax. In more details, comparisons with ground-based measurements indicate bias and dispersion with respect to validation data typically on the order or below 0.2 DU. Over clean areas, long-term average maps show large-scale systematic biases that are difficult to correct and can mask real (weak) signals. This was the actual motivation to develop the TROPOMI COBRA SO_2 product. As will be shown in the following chapters, much of the systematic offsets present in the operational product are solved with COBRA. The systematic VCD uncertainty (contribution from the COBRA spectral fit only) is very small, typically less than 0.04 DU. The SO_2 vertical column precision for individual pixel is in the range 0.5 - 1 DU, i.e., well within the requirement.

Generally speaking, the validation of a satellite SO_2 column product is a challenge for several reasons, on top of which is the representativeness of the correlative data when compared to the satellite retrievals. Another reason comes from relatively high noise of the satellite data (even with improved algorithm as COBRA). For anthropogenic SO_2, a very important limitation comes from the fact that the number of available correlative datasets in strongly polluted regions is very rare. Therefore, the evaluation of satellite SO_2 products is mainly based on satellite to satellite comparisons/consistency checks.

In this document, we will evaluate the SO2CBR product against existing TROPOMI SO_2 products and ground-based measurements from MAX-DOAS and direct sun instruments, at few stations.

5 Slant column verification

As a first step of the evaluation, it is necessary to compare the results from the COBRA SO_2 prototype algorithm as presented in [RD01] to the data produced by the algorithm implemented in PAL. Note that here only the SO_2 slant column is investigated, as the COBRA approach deals only with the spectral fitting part of the algorithm (the AMF and error modules are very close to the operational implementation). Figure 1 illustrates the status of the implementation for one week of TROPOMI test data, as an example for the slant columns retrievals. The top panel corresponds to the averaged SO_2 VCD map from the prototype algorithm (labelled BIRA), the center panel is the same PAL implementation, and the lowest panel corresponds to SO_2 VCDs difference of PAL minus BIRA. A nearly perfect agreement is found between SCD results over the entire global domain. A similar match between prototype algorithm and PAL is found for other days/periods taken randomly.

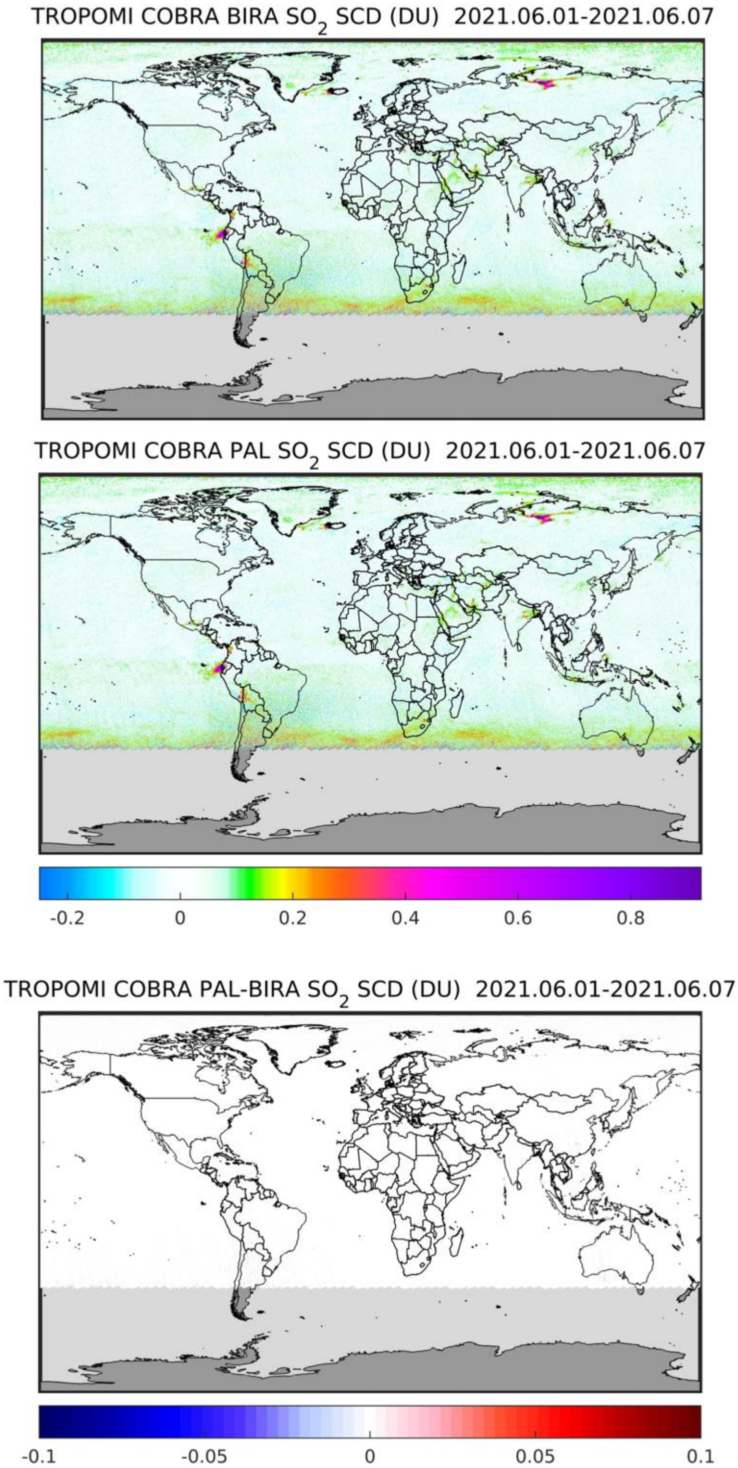

Figure 1: Comparison of SO_2 slant column data for one week (1-7 June 2021) for the prototype algorithm-BIRA (top) and the PAL implemented algorithm (center). The SO_2 VCDs difference map of PAL minus BIRA is show on the lowest panel.

6 Validation of SO2CBR vertical column density product

In the sections below, we evaluate the SO2CBR vertical column density (VCD) product through comparison with data from other satellite products and from ground based measurements.

6.1 Comparison with TROPOMI SO_2 products

Figure 2 compares the operational DOAS and COBRA seasonal averaged SO_2 VCD maps from September to November 2019. The data are gridded at a resolution of 0.1° x 0.1° and smoothed by a 2-dimensional 5-points box car function. Both DOAS and COBRA results are extracted using identical pixel selection criteria: SZA less than 60°, radiometric cloud fraction lower than 30% and TROPOMI rows 26-424. From Fig. 2, several artefacts are evident in the DOAS product. Negative values are found in the tropics and a large scale positive bias at mid-latitudes. In comparison, COBRA remarkably solves all the systematic biases found in the operational product whereas the signal from major SO_2 sources (e.g. in China, India, Middle-East, South Africa, Central and South America) is nicely preserved. Note that for individual pixels with unambiguous detection of SO_2 (typically SO_2 VCDs larger than 2 DU), the agreement between DOAS and COBRA is excellent (not shown).

Retrieval results using COBRA are also evaluated in Fig. 2 against a scientific TROPOMI SO_2 product generated using the PCA approach (see [RD01] for details). As can be seen, an overall excellent agreement is found between COBRA and PCA retrievals, the observed SO_2 spatial distributions being essentially the same. When comparing the TROPOMI COBRA and PCA maps, very consistent results are found. Yet, the quality of COBRA seems slightly better than the PCA retrievals. In particular, COBRA is much less sensitive to the South Atlantic Anomaly than PCA data, which exhibit many outliers in the corresponding region. At mid-latitudes, there is also a slight positive bias (of about +0.1 DU on average) and higher noise in the PCA results compared to COBRA.

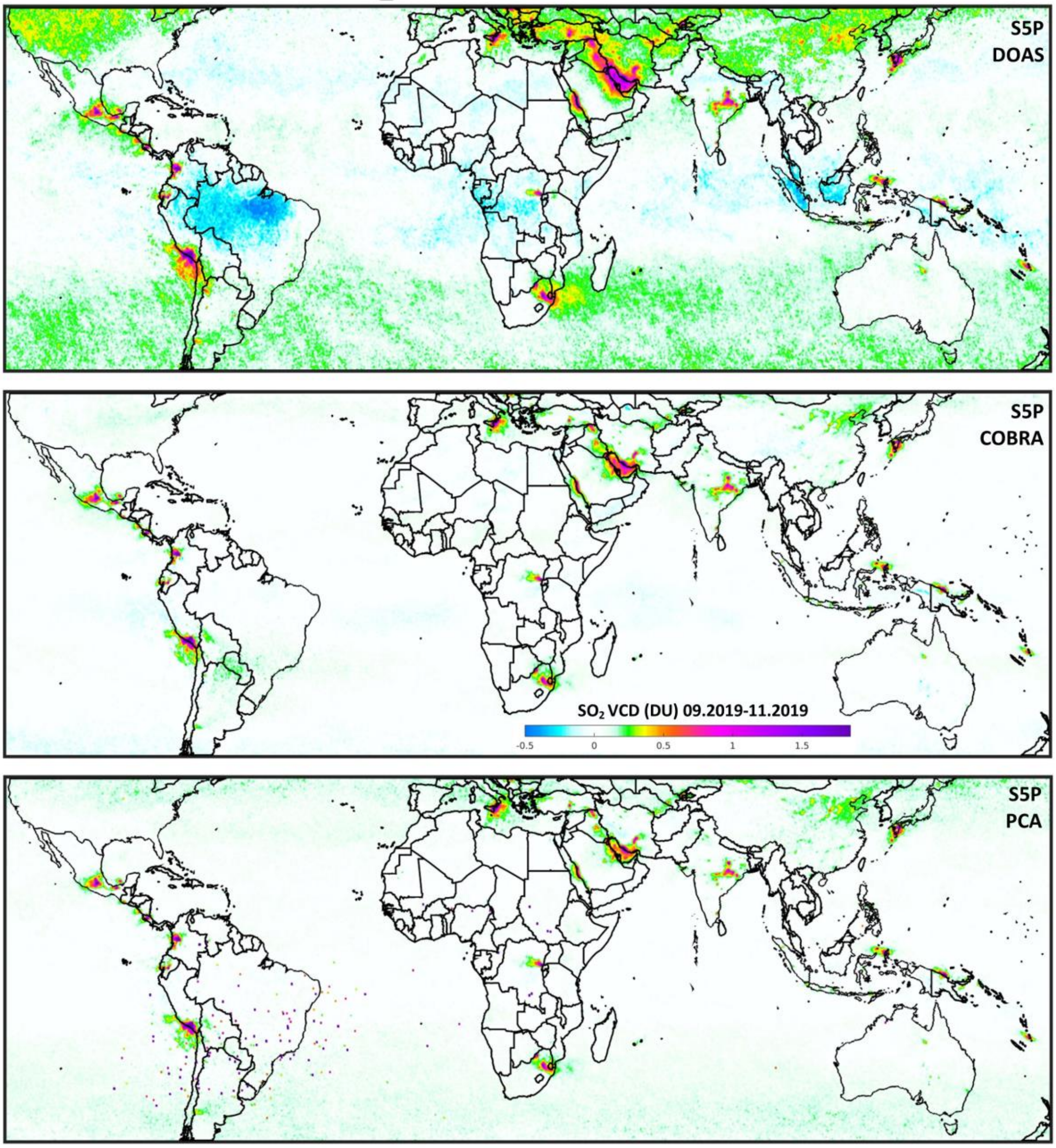

Figure 2: Comparison of seasonal mean SO_2 columns for September to November 2019 retrieved from TROPOMI DOAS, COBRA and PCA algorithms (from top to bottom). Consistent pixel selection criteria, gridding and retrieval settings are applied. For all four data sets, a fixed AMF of 0.4 is applied.

6.2 Comparison with ground-based data.

6.2.1 TROPOMI COBRA versus MAX-DOAS measurements

Here we compare our TROPOMI SO_2 VCD data to Multi-Axis DOAS (MAX-DOAS) observations at two sites, both characterized by relatively low SO_2 columns: Xianghe (China, 39.77°N, 117°E) and Mohali (India, 30.67°N, 76.74°E). The reader is referred to [RD01] for further information on the MAX-DOAS instrument and retrievals. For the comparison, we have used a common set of selection criteria for the satellite data. For each day, we selected the TROPOMI pixels within a 25 km radius circle around the station of interest, a strict radiometric cloud fraction threshold of 20%, SZA lower than 60° and AMF larger than 0.2. If the number of retained pixels is larger than 10 then the mean SO_2 VCD is calculated and compared to the averaged SO_2 column for the MAX-DOAS measurements within ±1 hour of the S-5P overpass time.

The comparison results between COBRA and MAX-DOAS measurements are shown in Figure 3a and 3b, for Xianghe and Mohali, respectively. Overall, the agreement between COBRA and MAX-DOAS data is very good, keeping in mind that the levels of SO_2 columns are quite low. The slopes of the regression lines are close to unity. The mean SO_2 columns from MAX-DOAS and TROPOMI COBRA retrievals are comparable and of ~ 0.2 DU at each station for SZAs less than 50° (slightly higher at large angles). This further supports the idea that COBRA is capable to reproduce low SO_2 column results.

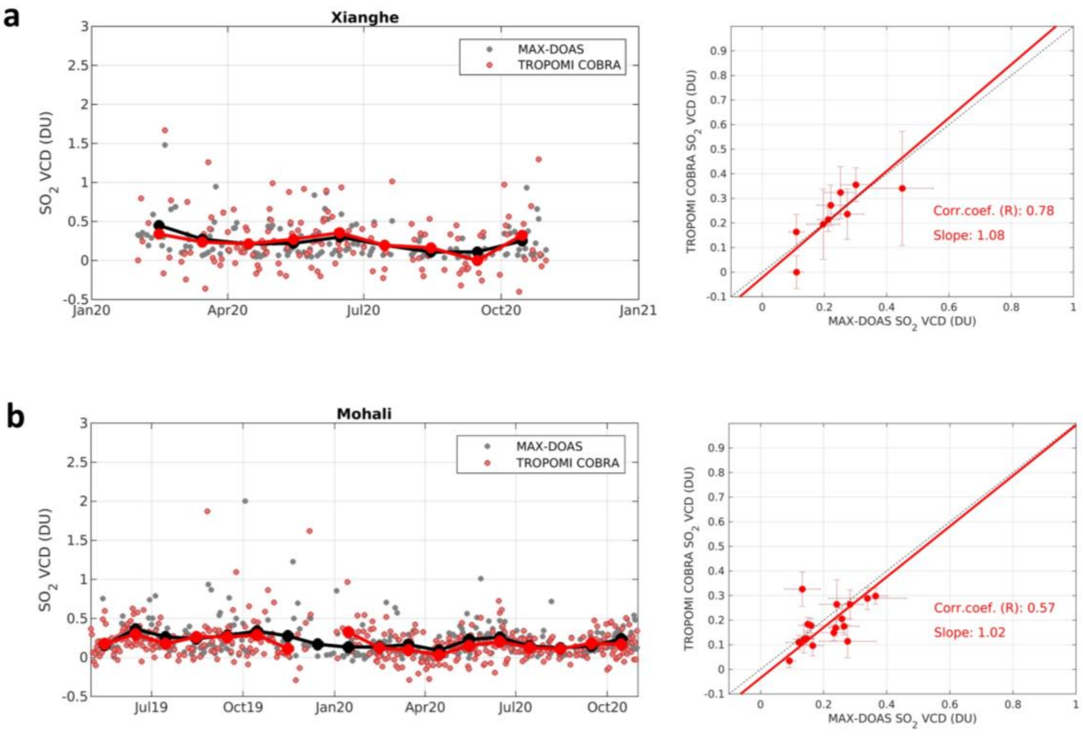

Figure 3: (left) Comparison of monthly mean SO_2 columns from MAX-DOAS and TROPOMI COBRA for (a) Xianghe, and (b) Mohali. The grey and pale red dots correspond to the individual days. (right) Scatter plots of monthly mean SO_2 columns of TROPOMI COBRA vs MAX-DOAS observations. Error bars are the standard errors on the monthly average SO_2 columns. The correlation coefficient and slope of the regression line are given as an inset for each plot.

6.2.2 TROPOMI COBRA versus Pandora measurements

Figure 4: The ground measurements stations for the Pandonia Global Network. Image from PGN website [URL03], last access 30-06-2022.

In order to validate the SO_2 vertical columns from the SO2CBR product, ground-based measurements from the Pandonia Global Network have been investigated. The PGN covers almost the full geographical latitude range. The stations are visualized in [URL03]. Although many of these locations are well suited for ground based atmospheric measurements of trace gases such as tropospheric NO_2, only a handful of locations are located in SO_2 source areas and are thus meaningful for comparison with the SO_2 product (see geographical distribution of SO_2 emission hotspots detected by COBRA at [URL02]). Moreover, at the time of writing, the retrieval of SO_2 from Pandora direct sun measurements is performed in a demonstration phase and not yet for all stations. Because of this, only two stations from PGN have been used for the comparison: MexicoCity-Vallejo (Mexico, 19.48°N, 99.14°W, #157) and Wakkerstroom (South Africa, 27.34°S, 30.14°E, #159). In addition to the data availability, these two stations were selected because of the relatively large measured SO_2 VCDs. In the future, it is anticipated that SO_2 columns from more PGN stations will become routinely available and could be used for the validation of SO2CBR.

Figure 5 compares the TROPOMI SO2CBR results with PGN SO_2 VCD values at the 2 stations from September 2019 to August 2021. All cloud-free TROPOMI pixels (with cloud fraction<20%) falling within 25 km from the station are considered. The mean (and std) SO_2 VCDs (for the polluted scenario) are calculated for each day and compared to averaged and maximum column values from the Pandora measurements within +/- 1 h aroud TROPOMI overpass. Overall, TROPOMI and the ground-based data agree reasonably well with each other. The regression analysis reveals a favorable correlation coefficient between satellite and ground-based column estimates of 0.76 at both sites. In terms of slopes, it is of 1.78 at Mexico City and 0.78 at Wakkerstroom. It should be stressed, however, that the SO_2 columns are generally rather low (often <1 DU) and that the slopes estimates are to some extend influenced largely by a few data points with large columns. The situation at Mexico City is particularly difficult with multiple SO_2 sources (of volcanic and anthropogenic origins) contributing to the observed SO_2 VCD at different altitudes. More work would be needed to understand the differences. Finally, we note that similar comparison using the TROPOMI operational SO_2 product has been carried out as part of the official S5P operational validation [URL04, RD03].

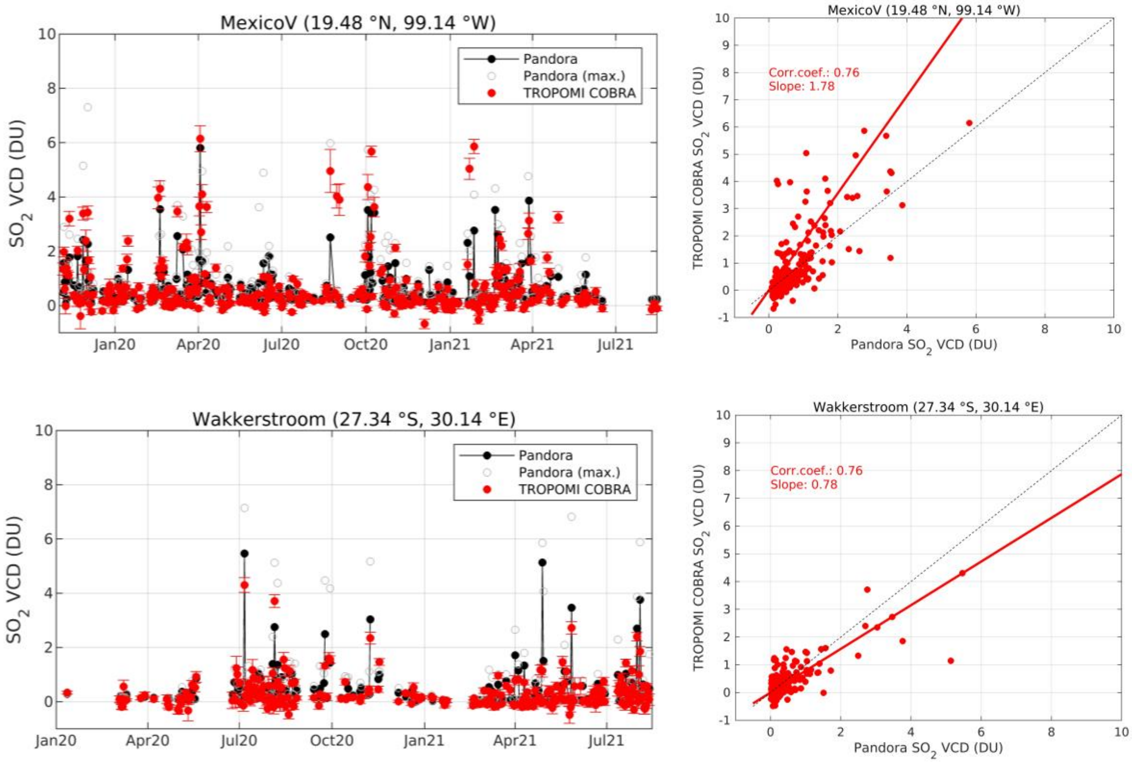

Figure 5: Comparison of TROPOMI COBRA SO2 VCDs to preliminary Pandora measurements (daily means) at Mexico City and Wakkerstroom from September 2019 to August 2021. The following selection criteria were applied: distance < 25km, CF<0.2, time window +/- 1h around overpass.

7 Improvement of the SO2CBR v2 product quality

Inspecting the retrievals from SO2CBR version1, it turns out that the results are often of bad quality in case of a strong eruption (see Figure 6, top panel). The root cause of this artefact turns out to be related to the initial covariance matrix (used for the retrieval of SO_2 slant columns) that can be strongly affected by spectra containing SO_2 absorption. In principle, this is mitigated by excluding the corresponding spectra from the covariance calculations for the next iterations. However, this procedure was found far than perfect and there are situations where the product quality was degraded for part of the orbit, often leading to negative biases. In the SO2CBR version 2, this has been largely improved by using DOAS retrieved slant column densities (from the operational product) to prefilter the spectra for the initial covariance matrix calculation. Most of the related artefacts are solved by this procedure as demonstrated in Figure 6 (bottom panel)

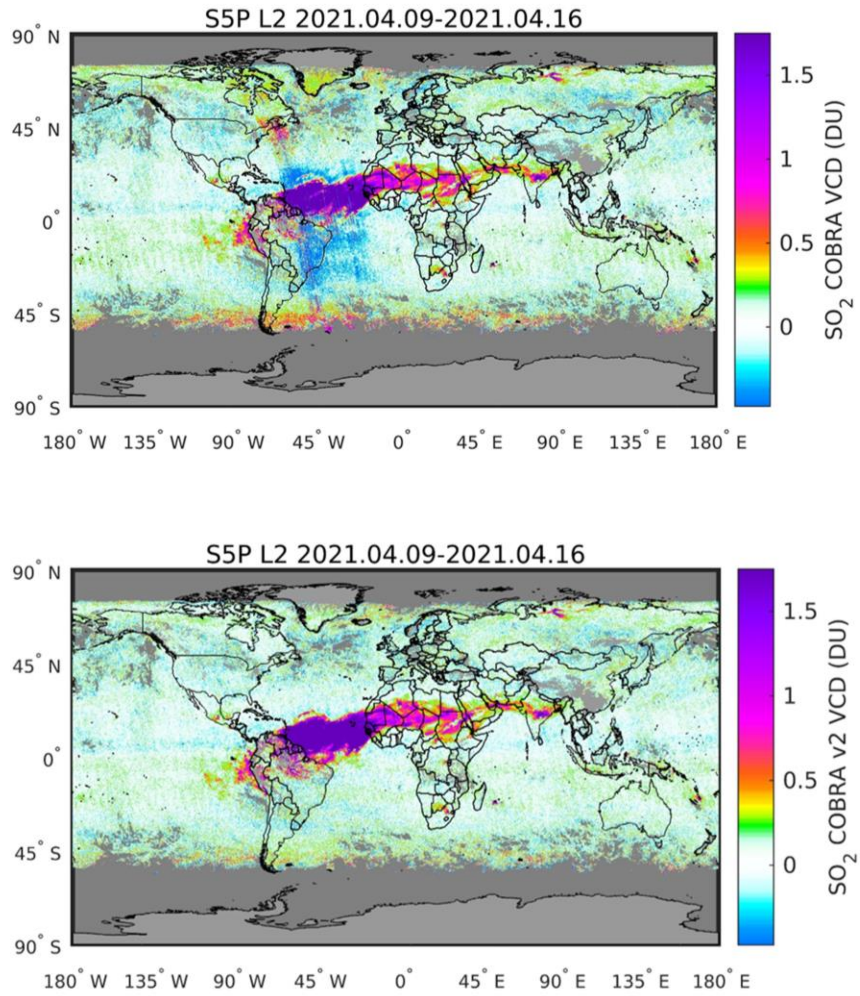

Figure 6: Comparison of TROPOMI COBRA SO2 VCDs from (top) version 1 and (bottom) version 2, averaged over the period from 9 to 16 April 2021. The large SO2 plume observed in the tropical band is from the eruption of La Soufrière.

Another improvement of SO2CBR version 2, compared to version 1, is for high latitudes. Owing to a better slant column error formulation, the selection of spectra for the covariance matrix calculation is improved at high latitudes. The effect on the retrievals can be appreciated on Figure 7 for the complete months of data. In SO2CBR version 1, there is a general positive/negative bias at high latitudes and over certain regions (e.g., Greenland), as shown on top panels of Figure 7. These artefacts are largely reduced in SO2CBR version 1 (Figure 7, bottom panels).

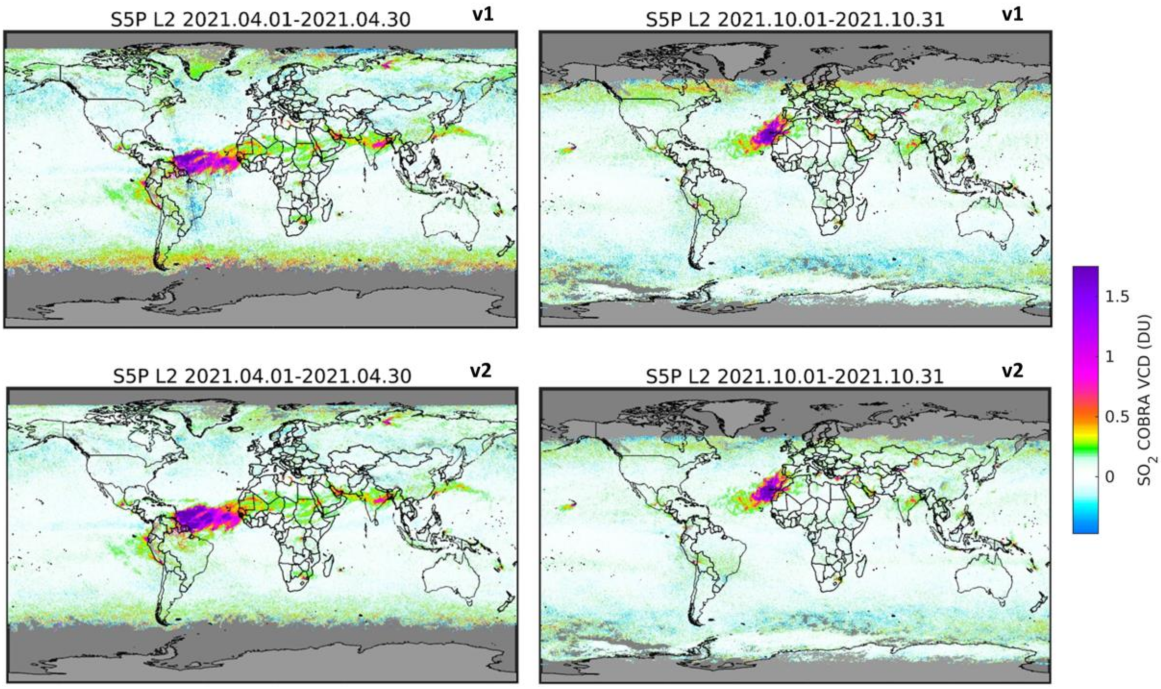

Figure 7: Comparison of TROPOMI SO2 VCD monthly averages (left: April 2021, right October 2021) from COBRA version 1 (top) and COBRA version 2 (bottom). The extended SO2 plume in October 2021 near the West coast of Africa is due to the eruption of Cumbre Vieja.

8 Conclusions

This document described the validation exercise of data from the S5P COBRA SO_2 column product (SO2CBR, version 02.00.00).

The SO_2 SCDs compare very well with retrieved SCDs obtained from the prototype COBRA algorithm, illustrating that the implementation of the code in PAL is a success.

Compared to the TROPOMI SO_2 products from the operational and scientific PCA algorithms, the results from SO2CBR are improved over clean areas and at high latitudes. For large emission hotspots, all products are consistent.

The agreement of SO2CBR VCD data with MAX-DOAS (Xianghe and Mohali stations) and Pandora (Mexico City and Wakkerstroom) instruments is generally very good. No discernible biases can be identified from the comparison.

For specific conditions, namely largely volcanic eruptions, and observations at high latitudes, the results from the SO2CBR version 02.00.00 are found better than SO2CBR version 01.00.00.