Copernicus Atmospheric Mission Performance Cluster Service

Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products #15: April 2018 – May 2022

Prepared by: Copernicus Atmospheric Mission Performance Cluster Service
Reference: SSP-MPC-IASB-ROCVR-15.01.00-20220713
CI identification: DI-MPC-ROCVR
Document update: #15
Issue: 15.01.00
Date of issue: 13 July 2022
Status: Final
Distribution: Public
Disclaimer

The information provided in this document has been produced in the context of the Copernicus Atmospheric Mission Performance Cluster Service (ATM-MPC). The activities leading to those results have been contracted by the European Space Agency (ESA) and coordinated by the Royal Belgian Institute for Space Aeronomy (BIRA-IASB) under supervision of the Royal Netherlands Meteorological Institute (KNMI) as the prime contractor. All information in this document is provided “as is” and no guarantee or warranty is given that the information is fit for any particular purpose. For the avoidance of all doubts, ESA, BIRA-IASB and KNMI expressly disclaim any and all warranties, express or implied, including without limitation warranties of merchantability and fitness for a particular purpose, with respect to the information. In no event shall ESA, BIRA-IASB and KNMI be liable for any direct, indirect, incidental, punitive, or consequential damages of any kind whatsoever with respect to the information. The user thereof uses the information at its sole risk and liability.
## Approval Record

<table>
<thead>
<tr>
<th>Checked by</th>
<th>Approved by</th>
<th>Signatures</th>
</tr>
</thead>
<tbody>
<tr>
<td>J.-C. Lambert (BIRA-IASB)</td>
<td>A. Dehn (ESA)</td>
<td>J.-C. Lambert (BIRA-IASB)</td>
</tr>
<tr>
<td>Q. L. Kleipool (KNMI)</td>
<td></td>
<td>[Signature]</td>
</tr>
<tr>
<td>D. Loyola (DLR)</td>
<td></td>
<td>[Signature]</td>
</tr>
<tr>
<td>J. P. Veefkind (KNMI)</td>
<td></td>
<td>[Signature]</td>
</tr>
<tr>
<td>A. Dehn (ESA)</td>
<td></td>
<td>A. Dehn (ESA)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Signature]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Digitally signed by Jean-Christopher Lambert (Signature)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Date: 2022.07.19 09:36:44 +02'00'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Angelika Dehn</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Digitally signed by Angelika Dehn</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Date: 2022.07.19 09:58:20 +02'00'</td>
</tr>
</tbody>
</table>
# Document Information

## Document Identification

<table>
<thead>
<tr>
<th>Title</th>
<th>Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products #15: April 2018 – May 2022</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of document</td>
<td>ATM-MPC S5P Routine Operations Consolidated Validation Report (ROCVR)</td>
</tr>
<tr>
<td>Document ID</td>
<td>S5P-MPC-IASB-ROCVR-15.01.00-20220713</td>
</tr>
<tr>
<td>ROCVR update Issue number</td>
<td>#15 version 15.01.00</td>
</tr>
<tr>
<td>Date of issue</td>
<td>13 July 2022</td>
</tr>
<tr>
<td>Status</td>
<td>Final</td>
</tr>
<tr>
<td>Distribution</td>
<td>Public</td>
</tr>
<tr>
<td>Available on:</td>
<td><a href="https://mpc-vdaf.tropomi.eu">https://mpc-vdaf.tropomi.eu</a></td>
</tr>
<tr>
<td>Editors</td>
<td>A. Keppens and J.-C. Lambert (BIRA-IASB)</td>
</tr>
</tbody>
</table>
Executive Summary

This document reports consolidated results of the Routine Operations Validation Service for the Sentinel-5 Precursor (S5P) Tropospheric Monitoring Instrument (TROPOMI) \[ER_TROPOMI\]. S5P TROPOMI contributes to the space component of the European Earth Observation programme Copernicus \[ER_CoperESA\]. The S5P Routine Operations Validation Service is provided by the Atmospheric Mission Performance Cluster (ATM-MPC) for Level-1 and Level-2 data products generated by the Near Real Time (NRTI) and Offline (OFFL) processors since the first public data release in July 2018. The present Routine Operations Consolidated Validation Report (ROCVR) integrates results from the MPC Validation Data Analysis Facility (VDAF) \[ER_VDAF\] with ad hoc support from S5P Validation Team (S5PVT). The S5P Routine Operations Validation Service details and complements the conclusions and features described in the Product Readme Files (PRF) delivered with the S5P products, in which users can find practical recommendations on S5P data usage. The present report covers the period of S5P operation from April 2018 until May 2022. It includes validation results for version 02.02.00 and 02.03.01 of the NL-L2 processor suite (NO2, CO, CH4, AAI, and ALH data), version 02.03.00 of the UPAS processor suite (total and tropospheric O3 columns, HCHO, SO2, and CLOUD data), and validation results for the operational ozone profile data (processor NL-L2 v02.03.01).

Radiance and Irradiance

The validation of the wavelength assignment of the S5P L1B_UVN v02.00.00 products concludes to an agreement of within 0.01 nm, which is within the pre-launch calibration uncertainty. The reflectance in bands 1-3 is 1 % to 3 % lower than TOMS and the used ice radiance model. The radiance in bands 1-3 is up to 5 % smaller than OMPS-nadir radiance, above 320 nm this is a wavelength independent bias. Below 320 nm the wavelength dependence seems to vary with the latitude. In band 1 around 280 nm the radiance deviates more than 10 % from OMPS values. The absolute radiometric calibration for UV radiance lacks accuracy and as a result may be updated in the future. In the spectrally overlapping regions of bands 2 and 3 there is a discrepancy of about 2 % in the L1b radiance signals. The radiance in band 6 was compared to model spectra in the continuum around the O2A band. The signal of TROPOMI is 1-2% lower than the model. For bands 1 to 6 (UV, UVIS and NIR) degradation has been observed for the radiance. The degradation is the largest at short wavelengths. The decrease in radiance signal per 1000 orbits is between 0.31% in band 1 and 0.02% in band 6. The degradation is planned to be corrected in a future update of the calibration key data. The absolute and relative radiance radiometry of the SWIR bands were validated using reference stations in Railroad Valley and in the Saharan desert. Current validation results give upper limits of <5% for the absolute calibration and <0.8% for the relative calibration. The absolute irradiance calibration of TROPOMI has been compared to other published solar reference datasets. After an update to the calibration based on OMPS-nadir data, the UV and UVIS spectrometers agree within 0-5 % with the references. For extreme swath angles the deviations are larger in the UV. For the NIR spectrometer the irradiance spectrum is approximately 1.5 %-3.5 % lower than the reference spectra. The SWIR spectrum is approximately 0.6 % lower than the closest reference spectrum.
Ozone Column

The S5P L2_O3 NRTI and OFFL total ozone column data are in good overall agreement with correlative ground-based measurements from the Brewer, Dobson and NDACC ZSL-DOAS/SAOZ monitoring networks, and with the Metop-B GOME-2, Aura OMI, and Suomi-NPP OMPS nadir satellite instruments. Across the networks the mean bias of about +0.8 % (NRTI) and +0.3 % (OFFL) and the standard deviation of the relative difference both comply with mission requirements, that is, a bias lower than 5 % and an uncertainty due to random errors (dispersion) better than ±2.5 %. The instrumental switch to smaller (along-track) ground pixels on the 6th of August 2019 did not affect the agreement with the ground-based reference data. The upgrade of the UPAS Level-1b-to-2 processor to version 2.1.3 on 16 July 2020 (NRTI) and 13 July 2020 (OFFL), and to version 2.1.4 in November 2020, improves the agreement of the NRTI product with both the OFFL product and with ground-based measurements at stations with complex surface albedo (snow/ice). Analysis of a diagnostic data set (DDS) suggested that the upgrade to version 2.2.1 in early July 2021, using v2 L1b data, caused a slight increase in the S5P ozone columns, of approximately 0.7 % for NRTI and 1.5 % for OFFL. Based on the ground-based comparisons available since July 2021, no significant bias change is observed for the NRTI product, and an increase of approximately 1 % is observed for the OFFL product. The recent upgrade to UPAS version 2.3.0 in March 2022 does not appear to affect the ground-based comparison results.

The comparison of S5P TROPOMI total ozone column data with other satellite data sets (GOME-2B, OMI, OMPS) over cloudy scenes highlights differences in the cloud models used in the retrieval algorithms. Large and/or systematic differences between satellite datasets also exist at high solar zenith angles (hence at high latitudes), and in the case of uncertain ground albedo. A very minor dependence (<1%) on across-track scan position was derived from analysis departures when assimilating the NRTI product in the CAMS system.

Tropospheric Ozone Column

The S5P L2_O3_TCL OFFL tropospheric ozone column data (CCD algorithm) are in good general agreement with correlative measurements from the ozonesonde monitoring network and from the Metop-B GOME-2 and Aura OMI satellite instruments. Across the ground-based network the mean bias (around +17 % or +3.5 DU) and the mean dispersion of the differences (about 26 % or 4.8 DU) both comply with mission requirements, that is, a bias lower than 25% and an uncertainty (dispersion) less than 25%. The Level-1 data processor was upgraded in July 2021. The comparison of updated S5P O3_TCL data to ozonesonde measurements at four stations indicates a slight change of +0.9±2.1 DU in bias between Level-1 versions 01.x and 02.x. Longer time series will be needed to assess whether subtler changes occur.

The bias between S5P and other data sets exhibits a seasonal cycle, which is most easily visible from satellite intercomparisons. A pattern of more elevated positive biases (7-10 DU or 25-60%) during the biomass burning season emerges at stations around the Atlantic equatorial basin. It is not clear whether this relation is causal. The interplay of satellite orbit and cloud coverage leads to two types of sampling error of up to 1 DU and ~5 DU, correlated in time and space (latitudinal stripes, patterns progressing along satellite orbit). Users can reduce sampling uncertainty by lowering the sampling resolution.
**Ozone Vertical Profile**

Comparison of the S5P L2_O3_PR ozone profile data (both NRTI and OFFL channels of the operational NL-L2 processor v02.03.01) with ozonesonde and lidar measurements concludes to a median agreement better than 10 % in the troposphere and up to the upper troposphere/lower stratosphere (UTLS), and to higher values (20 %) of dispersion. In the stratosphere comparisons reveal vertically oscillating biases of 5-10 % (positive and negative), with smaller dispersion than in the troposphere (order of 10 %). Chi-square tests demonstrate that on average the observed differences confirm the ex-ante satellite and ground uncertainty estimates in the stratosphere, while around the tropopause and below (around 15-20 km and lower), the mean chi-square value increases up to about two. Here, the predicted (random) satellite uncertainty is smaller than what is actually observed. The information content of the ozone profile retrieval is characterised by about six vertical sub-columns of independent information (estimated from the Degree of Freedom of the System) and a vertical sensitivity nearly equal to unity at altitudes from about 20 km (UTLS) to 50 km, and decreasing rapidly at altitudes above and below. The altitude registration of the retrieved profile information (estimated as the altitude of the barycentre of the associated averaging kernels) usually is close to the nominal retrieval altitude in the 20-50 km altitude range, and shows positive and negative offsets of up to 10 km below and above the 20-50 km altitude range, respectively. The effective vertical resolution of the profile retrieval (estimated as the Full Width at Half Maximum (FWHM) of the averaging kernels) usually ranges within 10-15 km, with a minimum close to 7 km in the middle stratosphere. Increased sensitivities and higher effective vertical resolutions (FWHM) can be observed for high solar zenith angles, as can be expected. On the other hand, a reduced DFS, sensitivity, and retrieval quality are sometimes observed for scenes that have both high SZA and high surface albedo, especially over the sea around the Antarctic. A detailed assessment of the effect of influence quantities will be possible in the future when more operational S5P ozone profile data become available.

Global maps of the six ozone profile sub-columns show the geographical coverage of the data, but also a slight along-orbit striping, especially in the middle stratosphere (24-32 km sub-column). Global maps of the integrated L2_O3_PR ozone profile data and of the L2_O3 total column data look mutually consistent. On the other hand, orbit curtain plots reveal that for some ground pixels the retrieved ozone profile deviates strongly and non-physically from the a-priori profile used in the retrieval. This issue needs further examination. Finally, in the absence of clouds, data files sometimes contain negative surface albedo values. The ground pixels affected by this anomaly are usually located at the east and west edges of the TROPOMI measurement swath.

**Nitrogen Dioxide**

The three S5P L2_NO2 data products (tropospheric, stratospheric, and total column) up to version 02.03.01 (NRTI, OFFL) and 01.02.02 (RPRO) are in good overall agreement with correlative ground-based measurements. This assessment covers the S5P operation period from April 2018 until May 2022. Correlative measurements from the MAX-DOAS network (troposphere), the NDACC ZSL-DOAS/SAOZ network (stratosphere), and the Pandonia Global Network (total), respectively, as well as correlative satellite data products (OMI), are used as validation references. Generally, the bias between S5P and ground-based data increases with NO2 column amount except for stratospheric columns. Similar biases and uncertainty estimates are detected for the L2_NO2 NRTI and OFFL/RPRO datasets using ground-based validation.
The L2_NO2 tropospheric column data is compared to MAX-DOAS ground-based data from 27 stations. A negative median bias of -34% is found. This bias depends on pollution level at the station, being positive (15%) over clean areas (<2 Pmolec/cm²) and negative (-46%) over highly polluted areas (>15 Pmolec/cm²). The tropospheric bias is within the mission requirement of 50%. This bias estimate can be reduced by up to 20% when MAX-DOAS profile data are vertically smoothed using the SSP averaging kernels. The median dispersion of about 2.6 Pmolec/cm² exceeds the mission precision requirements (0.7 Pmolec/cm²). It is only within limits for very clean stations. Use of the reprocessed PAL data for VDAF-AVS comparisons show a reduction of bias from -26% to -16%.

Comparison of TROPOMI tropospheric NO2 column data to OMI measurements concludes to differences of more than -20% in winter time over polluted regions (China, Europe) and up to +20% in the clean Pacific region. The bias decreases when only V1.4 and later ones are used for comparisons.

The NRTI L2_NO2 stratospheric column data is compared to ZSL-DOAS UV-visible ground-based measurements at 25 NDACC stations distributed from pole to pole. Accounting for the large horizontal smoothing of stratospheric NO2 in the zenith-scattered-light (ZSL) geometry and for the NO2 diurnal cycle, SSP reports stratospheric NO2 column values generally lower by approximately 0.17 Pmolec/cm². The median bias of -6% is within the SSP mission requirements (0.2-0.4 Pmolec/cm²) as well as the dispersion of 0.3 Pmolec/cm², considering the combined random errors and irreducible co-location mismatches. The OFFL L2_NO2 bias in comparison to the VDAF-AVS subset (13 stations) is -4.6% and the dispersion 0.3 Pmolec/cm². The comparison of OFFL L2_NO2 stratospheric columns with ground-based NDACC FTIR data shows a positive median bias for 24 stations of +4.7% with a dispersion of 0.3 Pmolec/cm². Even larger biases are observed at high-latitude (9 to 14%) and tropical stations (11 to 12%). Note that the SSP NO2 stratospheric data are averaged within 50km around the location of the best-estimated collocation, and the FTIR data are averaged within +/-1h around the SSP overpass time. Using PGN total NO2 column data at 3 high altitude stations, a bias of approximatively -10% is detected.

The L2_NO2 total NO2 column data are compared to ground-based Pandora column data at 36 PGN stations. The median bias between SSP and PGN data is -6.7% with a dispersion of 1.5 Pmolec/cm². Comparison results vary with the total amount of NO2, having a positive bias of 4% over clean stations and high mountain areas (< 6 Pmolec/cm²) and a negative bias of about -21% over polluted stations.

The version upgrades 01.04.00-02.03.01 improved bias and dispersion for tropospheric and total columns. Comparing 3-month periods (January-March) of different years, the bias was reduced from about -38% in 2019 to -18% in 2022 for tropospheric NO2 and from -6% in 2019 to -4% in 2022 for total NO2. The negative stratospheric bias also improved slightly.

**Formaldehyde**

The SSP L2_HCHO column data up to version 02.03.00 (NRTI/OFFL) and 01.01.05 (RPRO) is in good overall agreement with independent ground-based measurements from the NDACC FTIR and MAX-DOAS monitoring networks and to corresponding Aura OMI satellite data. Ground-based validation concludes to similar bias and uncertainty (dispersion) estimates for the L2_HCHO NRTI and L2_HCHO OFFL/RPRO dataset.

Comparisons with FTIR data from 28 stations show a negative bias of -29% for high emission stations (> 8 Pmolec/cm²) and a positive bias of 28% for clean stations (< 2.5 Pmolec/cm²) when vertical smoothing differences are minimized by application of the averaging kernels. The bias with respect to MAX-DOAS HCHO measurements is slightly higher with about -37%. This bias reduces to -20% when also using averaging kernels. Also here biases differ for clean (+10%) and polluted stations (-27%). All biases are within the SSP mission requirements (40-80%).
The dispersion versus FTIR data at clean stations of about 8 Pmolec/cm² and 10 Pmolec/cm² versus MAX-DOAS is within the SSP uncertainty mission requirements of 12 Pmolec/cm². These values are based on the use of median deviations to reduce the influence of larger outliers and vertical smoothing. The bias in comparison to OMI is less than -10% for most regions with some larger negative biases in Europe, Northern America and China (<20%). The dispersion of differences is about 2 Pmolec/cm² when considering regionally averaged columns. The processor upgrades to version 02.01.04 in December 2020, to version 02.02.01 in July 2021, and to version 02.03.00 in March 2022 (improved background correction), had no effect on SSP L2_HCHO data quality.

**Sulphur Dioxide**

The S5P L2_SO2 (NRTI and OFFL) sulphur dioxide column data are found in general good agreement with ground-based MAX-DOAS and PGN measurements and with other satellite observations from OMI and S-NPP OMPS. The bias and dispersion with respect to validation data are typically below 0.2 DU. From these comparisons it can be concluded that over polluted regions the SSP mission requirements are fulfilled. Over volcanic plumes the requirement on the bias is fulfilled, 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. Validation of the UPAS processor upgrade to version 02.01.03 released in July 2020 (and subsequent version 02.01.04 released in November 2020) shows good consistency of the SO2 data product before and after the processor switch. The processor upgrades to version 02.01.04 in December 2020, to version 02.02.01 in July 2021, and to version 02.03.00 in March 2022 (improved background correction), had no effect on SSP L2_SO2 data quality.

**Carbon Monoxide**

The S5P L2_CO (NRTI or RPRO concatenated with OFFL) carbon monoxide total column data is in good overall agreement with correlative measurements from the NDACC, TCCON, and COCCON FTIR monitoring networks. It exhibits a positive bias of approximately +10 % (NRTI, before July 2019) or +6.5 % (OFFL and NRTI after July 2019) on an average, which falls well within the mission requirement (bias of maximum 15 %). The standard deviation of the relative bias is on an average 5 % against NDACC, TCCON, and COCCON, which is also within the mission requirement for precision (better than 10%). The averaged correlation coefficient reaches 0.9 for NDACC, TCCON, and COCCON. From July 3 2019, onwards the NRTI processor uses the same settings as the OFFL processor and both products perform similarly since then. The processor update to version 02.02.00 on July 1 2021 introduces a de-striped CO column product, which shows less scatter when compared to the ground-based FTIR columns. This processor update includes a change in spectroscopic parameters and preliminary results using rapid delivery NDACC data indicate that the bias is reduced to 2.9 %.

**Methane**

The S5P L2_CH4 (OFFL concatenated with RPRO) methane total column averaged data is in good overall agreement with correlative measurements from the NDACC, TCCON and COCCON FTIR monitoring networks. The standard and bias-corrected SSP xCH4 column data exhibit a negative bias against TCCON of -0.47 % and -0.03 % respectively, against COCCON of -0.66 % and -0.16 %, respectively and against NDACC of -0.5% and +0.3% respectively, which falls well within the mission requirement (bias of maximum 1.5 %). The standard deviation of the relative bias for TCCON and COCCON is on an average 0.6 % which is also within the mission requirement for precision (<1 %). The averaged correlation coefficient 0.70 is rather low, partly because not all outlying pixels are filtered with the qa_value above 0.5. Between March 11 2020 and July 1 2021, the number of pixels with qa_value above 0.5 is reduced due to a change in the cloud data used in the qa_value computation.
The implementation of the processor update with version 02.02.00 on July 1 2021, which uses an updated cloud fraction definition, improved spectroscopic parameters and a new a-posteriori bias correction independent of any reference data, will be evaluated in a future ROCVR when TCCON data is updated and more recent measurements are available. A preliminary comparison with rapid delivery NDACC data is included. The introduction of sun-glitz pixels on November 14 2021 with processor 02.03.01 is evaluated using the scientific product available at SRON and shows similar behaviour as the standard product. The validation against TCCON shows a mean bias of -0.55% for the bias-corrected product for the limited co-locations found. Comparison of the S5P scientific L2_CH4 product against GOSAT XCH4 over the ocean shows a mean bias of -0.24±0.85 % and a Pearson’s correlation coefficient of 0.8, and over land shows a mean bias of -0.7±0.8 % and a Pearson’s correlation coefficient of 0.87. The S5P scientific L2_CH4 data when compared to S5P WFMD-DOAS XCH4 product over ocean shows a mean bias of -0.14± 1.4 % and a Pearson’s correlation coefficient of 0.56.

**Clouds**

The S5P L2_CLOUD (NRTI and OFFL) CAL and CRB radiometric cloud fractions were compared to other satellite data. S5P L2_CLOUD, S5P FRESCO, Aura OMI OMCLDO2 and Aqua MODIS are found to capture similar meridian variations. The filter guideline of qa_value≥0.5 is too strict for some applications, but we found that the less strict qa_value≥0.25 alternative (see CLOUD PRF) does not change the overall picture on meridian variation. Comparing spatially co-located measurements, the CLOUD CRB cloud fraction is slightly below that of S5P FRESCO, with higher deviations over high-latitude and/or coastal stations. This difference is reduced with the upgrade of the UPAS CLOUD processor from version 1 to version 2, likely due to the improvement in surface albedo and quality flagging. This reduced difference persists with the upgrades of the UPAS CLOUD processor from version 02.01.03 to 02.03.00 and FRESCO from 01.04.00 to 02.03.01. Furthermore, the difference dispersion is reduced with the CLOUD and FRESCO upgrades in July 2021.

The S5P L2_CLOUD (NRTI and OFFL) CRB cloud height data and CAL cloud top height data were compared to ground-based measurements from the CLOUDNET and ARM networks, and with S5P FRESCO and other satellite products. Note that the sensitivity of the TROPOMI NIR observations to clouds differs significantly from the sensitivity of CLOUDNET/ARM lidar/radar instruments used as a reference, and the error associated with the reference observations is also not yet included in those comparisons. With the upgrade of the UPAS CLOUD processor from version 1 to 2 (July 2020), deviations at high-latitude and/or coastal stations are reduced (e.g., Ny-Ålesund, Andoya, Iquique, Mindelo, Punta-Arenas), CRB cloud height is reduced compared to FRESCO and the dispersion between CAL and FRESCO cloud height is reduced. The upgrades to CLOUD 02.02.01 and FRESCO 02.02.00 seem to further reduce this dispersion at some sites. Given their different nature, we look at low clouds and high clouds separately to derive representative quality indicators (distinction set at CLOUDNET cloud top height 4 km). For low clouds, the bias of CAL cloud top height vs. CLOUDNET cloud top height is within mission requirements, but not for high clouds. The bias of CRB v2 cloud height vs. CLOUDNET cloud mean height is beyond mission requirements and for v2 is slightly below the CLOUDNET cloud base height. The dispersion between SSP and other cloud height data exceeds the SSP mission requirements, but likely the comparison error (due to the different sensitivity) is an important cause of this apparent discrepancy. For low clouds, the correlation of L2_CLOUD CAL and CRB cloud height vs CLOUDNET is markedly better for v2 than for v1.

Two fictitious geographical patterns in cloud parameters were identified for UPAS processor version 01.0x.0x: increased values for cloud top height and cloud fraction at the right edge of the swath at certain latitudes, and a North-South gradient for cloud albedo. These patterns are reduced for the processor upgrade from version 02.01.03 to 02.02.01 in July 2021.
Aerosol Index

The SSP L2_AER_AI (NRTI and OFFL) UV Aerosol Absorbing Index is in good overall agreement with similar satellite data products from EOS-Aura OMI and Suomi-NPP OMPS. Although compliant with the mission requirement of 1 UVAI unit in 2018, the bias observed with version 1.04.00 was larger than 1 UVAI unit as compared to OMI and OMPS. Currently the 02.03.01 version is within mission requirements due to the recent update of the L1b data including a correction for observed degradation of the irradiance combined with an offset factor applied to the UVAI data product. The observed bias and how it changes over time as related to wavelength-dependent degradation and corrections is under investigation. Trend analysis of UVAI is not recommended until reprocessing data (2022) is available.

Aerosol Layer Height

The SSP L2_AER_LH (OFFL) data product shows a very good agreement with two other satellite aerosol layer height estimates, from MISR (stereoscopic imagery) and CALIOP (active lidar sensing of the aerosol vertical distribution). SSP TROPOMI AER_LH shows a systematic difference with MISR aerosol plume height of about 600 m (lower for TROPOMI). This is mostly due to the difference in the sensitivity of the instruments and the differences in the algorithms. A difference of about 500 m (lower for CALIOP) is expected from simulations, TROPOMI ALH being sensitive to the centroid aerosol layer height. For very thick plumes the difference between TROPOMI ALH and CALIOP layer height even decreases to only 50 m. This is well within the requirements of 100 hPa for the bias. From the comparison with Earlinet backscatter profiles, also a very good agreement (R=0.91 over land, 0.59 over ocean), but a slightly larger difference (about -1000m over land, about -1500m over ocean) is found. At least part of these differences can be attributed to the different sensitivities of the TROPOMI and Lidar ALH measurements. The comparisons to the Earlinet stations indicate that TROPOMI ALH retrievals over bright surfaces should be treated with care.

The SSP L2_ALH dispersion is large due to cloud contamination and surface effects. With rigorous cloud screening, 50 % of the pixels are already within 1 km of the CALIOP weighted extinction height. Accounting for the expected bias, this is within the requirements of 50 hPa. But this preliminary conclusion needs further investigation and confirmation.

A limitation of the SSP TROPOMI ALH product has become apparent following the severe bushfires in New South Wales during the 2019-2020 fire season, which produced very high altitude smoke plumes (altitude > 20 km). These heights were not anticipated and ALH values are limited to about 13 km altitude. An update to include these very high altitudes is not foreseen for the near future.

Because of the degradation of the UVAI, the applied UVAI filter for the ALH retrievals removed a large and increasing number of observations, which are in principle well suited for ALH retrievals. Therefore, in July 2021, the UVAI filter was replaced by a filter based on a cloud mask. Due to this change the number of processed measurements increased by more than an order of magnitude.
# Processing Baseline Identification

This document reports consolidated validation results for the following S5P TROPOMI data products:

<table>
<thead>
<tr>
<th>Product ID</th>
<th>Stream</th>
<th>Version</th>
<th>In operation from (orbit #, date)</th>
<th>In operation until (orbit #, date)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1B_RA1/2/.../8</td>
<td></td>
<td>01.00.00</td>
<td>2818, 2018-04-30</td>
<td>19257, 2021-07-01 current version</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.00.00</td>
<td>2818, 2018-04-30</td>
<td>19257, 2021-07-01 current version</td>
</tr>
<tr>
<td>L1B_IR_UVN/SIR</td>
<td></td>
<td>01.00.00</td>
<td>2955, 2018-05-09</td>
<td>3943, 2018-07-18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.02</td>
<td>4245, 2018-08-08</td>
<td>5930, 2018-12-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.05</td>
<td>5931, 2018-12-05</td>
<td>7631, 2019-04-04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.06</td>
<td>7631, 2019-04-04</td>
<td>7999, 2019-04-30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.07</td>
<td>7999, 2019-04-30</td>
<td>12482, 2020-03-11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.08</td>
<td>12482, 2020-03-11</td>
<td>14285, 2020-07-16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.03</td>
<td>14285, 2020-07-16</td>
<td>16256, 2020-12-02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.04</td>
<td>16256, 2020-12-02</td>
<td>19307, 2021-07-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.01</td>
<td>19308, 2021-07-05</td>
<td>22810, 2022-03-09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.00</td>
<td>22813, 2022-03-09</td>
<td>current version</td>
</tr>
<tr>
<td>L2_O3</td>
<td>OFFL</td>
<td>01.01.07</td>
<td>7907, 2019-04-23</td>
<td>12431, 2020-03-07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.08</td>
<td>12431, 2020-03-07</td>
<td>14238, 2020-07-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.03</td>
<td>14239, 2020-07-13</td>
<td>16212, 2020-11-29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.04</td>
<td>16213, 2020-11-29</td>
<td>19257, 2021-07-01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.01</td>
<td>19258, 2021-07-01</td>
<td>22767, 2022-03-06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.00</td>
<td>22768, 2022-03-06</td>
<td>current version</td>
</tr>
<tr>
<td></td>
<td>RPRO</td>
<td>01.01.07</td>
<td>7907, 2019-04-23</td>
<td>7906, 2019-04-23</td>
</tr>
<tr>
<td>L2_O3_TGL</td>
<td>OFFL</td>
<td>01.01.05</td>
<td>2824, 2018-04-30</td>
<td>7421, 2019-03-20</td>
</tr>
<tr>
<td>(CCD)</td>
<td></td>
<td>01.01.06</td>
<td>7435, 2019-03-21</td>
<td>7791, 2019-04-15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.07</td>
<td>7804, 2019-04-16</td>
<td>12273, 2020-02-25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.08</td>
<td>12288, 2020-02-26</td>
<td>14075, 2020-07-01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.03</td>
<td>14132, 2020-07-02</td>
<td>16090, 2020-11-17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.04</td>
<td>16104, 2020-11-18</td>
<td>19140, 2021-06-20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.01</td>
<td>19311, 2021-07-02</td>
<td>22645, 2022-02-22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.00</td>
<td>22659, 2022-02-23</td>
<td>current version</td>
</tr>
<tr>
<td>L2_O3_PR</td>
<td>NRTI</td>
<td>02.03.01</td>
<td>21325, 2021-11-24</td>
<td>current version</td>
</tr>
<tr>
<td></td>
<td>OFFL</td>
<td>02.03.01</td>
<td>21207, 2021-11-16</td>
<td>current version</td>
</tr>
<tr>
<td>L2_NO2</td>
<td>NRTI</td>
<td>01.00.01</td>
<td>2955, 2018-05-09</td>
<td>3364, 2018-06-07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.00.02</td>
<td>3745, 2018-07-04</td>
<td>3946, 2018-07-18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.00</td>
<td>3947, 2018-07-18</td>
<td>5333, 2018-07-24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.02.00</td>
<td>5336, 2018-10-24</td>
<td>5929, 2018-12-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.02.02</td>
<td>5931, 2018-12-05</td>
<td>7517, 2019-03-27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.00</td>
<td>7519, 2019-03-27</td>
<td>7999, 2019-03-30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.01</td>
<td>8000, 2019-04-30</td>
<td>8906, 2019-07-03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.02</td>
<td>8907, 2019-07-03</td>
<td>16256, 2020-12-02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.04.00</td>
<td>16259, 2020-12-02</td>
<td>19306, 2021-07-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.00</td>
<td>19308, 2021-07-05</td>
<td>21222, 2021-11-17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.01</td>
<td>21223, 2021-11-17</td>
<td>current version</td>
</tr>
<tr>
<td>RPRO</td>
<td></td>
<td>01.02.02</td>
<td>2836, 2018-05-01</td>
<td>5235, 2018-10-17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.02.00</td>
<td>5236, 2018-10-17</td>
<td>5832, 2018-11-28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.02.02</td>
<td>5840, 2018-11-29</td>
<td>7424, 2019-03-20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.00</td>
<td>7425, 2019-03-20</td>
<td>7906, 2019-04-23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.01</td>
<td>7907, 2019-04-23</td>
<td>8814, 2019-06-26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.02</td>
<td>8815, 2019-06-26</td>
<td>16212, 2020-11-29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.04.00</td>
<td>16213, 2020-11-29</td>
<td>19257, 2021-07-01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.00</td>
<td>19258, 2021-07-01</td>
<td>21187, 2021-11-14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.01</td>
<td>21188, 2021-11-14</td>
<td>current version</td>
</tr>
<tr>
<td>Product ID</td>
<td>Stream</td>
<td>Version</td>
<td>In operation from (orbit #, date)</td>
<td>In operation until (orbit #, date)</td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>---------</td>
<td>----------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>L2_HCHO</td>
<td>NRTI</td>
<td>01.00.00</td>
<td>2955, 2018-05-09</td>
<td>3943, 2018-07-18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.01</td>
<td>3947, 2018-07-18</td>
<td>4244, 2018-08-08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.02</td>
<td>4245, 2018-08-08</td>
<td>5929, 2019-12-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.05</td>
<td>5931, 2018-12-05</td>
<td>7631, 2019-04-04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.06</td>
<td>7632, 2019-04-04</td>
<td>7999, 2019-04-30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.07</td>
<td>8000, 2019-04-30</td>
<td>12482, 2020-03-11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.08</td>
<td>12483, 2020-03-11</td>
<td>14285, 2020-07-16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.03</td>
<td>14285, 2020-07-16</td>
<td>16256, 2020-12-02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.04</td>
<td>16259, 2020-12-02</td>
<td>19306, 2021-07-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.01</td>
<td>19308, 2021-07-05</td>
<td>22810, 2022-03-09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.00</td>
<td>22813, 2022-03-09</td>
<td>current version</td>
</tr>
<tr>
<td>L2_SO2</td>
<td>NRTI</td>
<td>01.00.00</td>
<td>2955, 2018-05-09</td>
<td>3943, 2018-07-17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.01</td>
<td>3953, 2018-07-18</td>
<td>4244, 2018-08-08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.02</td>
<td>4245, 2018-08-08</td>
<td>5929, 2019-12-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.07</td>
<td>5931, 2018-12-05</td>
<td>12482, 2020-03-11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.08</td>
<td>12482, 2020-03-11</td>
<td>14285, 2020-07-16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.03</td>
<td>14285, 2020-07-16</td>
<td>16256, 2020-12-02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.04</td>
<td>16213, 2020-11-29</td>
<td>19307, 2021-07-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.01</td>
<td>19308, 2021-07-05</td>
<td>22810, 2022-03-09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.00</td>
<td>22813, 2022-03-09</td>
<td>current version</td>
</tr>
<tr>
<td>L2_CO</td>
<td>NRTI</td>
<td>01.00.00</td>
<td>5336, 2018-10-24</td>
<td>5929, 2018-12-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.02.02</td>
<td>5931, 2018-12-05</td>
<td>7517, 2019-03-27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.00</td>
<td>7519, 2019-03-27</td>
<td>7999, 2019-04-30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.01</td>
<td>7999, 2019-04-30</td>
<td>8906, 2019-07-03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.02</td>
<td>8906, 2019-07-03</td>
<td>16256, 2020-12-02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.04.00</td>
<td>16259, 2020-12-02</td>
<td>19306, 2021-07-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.00</td>
<td>19308, 2021-07-05</td>
<td>21222, 2021-11-17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.01</td>
<td>21222, 2021-11-17</td>
<td>current version</td>
</tr>
<tr>
<td>L2_CH4</td>
<td>RPRO²</td>
<td>01.02.02</td>
<td>0657, 2017-11-28</td>
<td>5346, 2018-10-25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.01</td>
<td>2818, 2018-04-30</td>
<td>5832, 2018-11-28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.02</td>
<td>2463, 2018-04-04</td>
<td>2477, 2018-04-05</td>
</tr>
<tr>
<td>Product ID</td>
<td>Stream</td>
<td>Version</td>
<td>In operation from (orbit #, date)</td>
<td>In operation until (orbit #, date)</td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>---------</td>
<td>----------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OFFL</td>
<td>01.03.02</td>
<td>2463, 2018-04-04</td>
<td>2477, 2018-04-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.02.02</td>
<td>5833, 2018-11-28</td>
<td>7424, 2019-03-20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.00</td>
<td>7425, 2019-03-20</td>
<td>7906, 2019-04-23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.01</td>
<td>7907, 2019-04-23</td>
<td>8874, 2019-06-26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.02</td>
<td>8812, 2019-06-26</td>
<td>16212, 2020-11-29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.04.00</td>
<td>16213, 2020-11-29</td>
<td>19257, 2021-07-01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.00</td>
<td>19258, 2021-07-01</td>
<td>21187, 2021-11-14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.00</td>
<td>21188, 2021-11-14</td>
<td>current version</td>
</tr>
<tr>
<td></td>
<td>NRTI</td>
<td>01.01.01</td>
<td>3947, 2018-07-18</td>
<td>4244, 2018-08-08</td>
</tr>
<tr>
<td>L2_CLOUD</td>
<td></td>
<td>01.01.02</td>
<td>4245, 2018-08-08</td>
<td>5929, 2018-12-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.05</td>
<td>5931, 2018-12-05</td>
<td>7631, 2019-04-04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.06</td>
<td>7631, 2019-04-04</td>
<td>7999, 2019-04-30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.07</td>
<td>7999, 2019-04-30</td>
<td>12482, 2020-03-11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.08</td>
<td>12482, 2020-03-11</td>
<td>14285, 2020-07-16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.03</td>
<td>14285, 2020-07-16</td>
<td>16256, 2020-12-02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.04</td>
<td>16259, 2020-12-02</td>
<td>19306, 2021-07-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.01</td>
<td>19308, 2021-07-05</td>
<td>22810, 2022-03-09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.00</td>
<td>22813, 2022-03-09</td>
<td>current version</td>
</tr>
<tr>
<td></td>
<td>OFFL</td>
<td>01.01.07</td>
<td>7907, 2019-04-23</td>
<td>12431, 2020-03-07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.01.08</td>
<td>12432, 2020-03-07</td>
<td>14238, 2020-07-13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.03</td>
<td>14239, 2020-07-13</td>
<td>16212, 2020-11-29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.01.04</td>
<td>16213, 2020-11-29</td>
<td>19257, 2021-07-01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.01</td>
<td>19258, 2021-07-01</td>
<td>22767, 2022-03-06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.00</td>
<td>22768, 2022-03-06</td>
<td>current version</td>
</tr>
<tr>
<td></td>
<td>RPRO</td>
<td>01.01.07</td>
<td>2818, 2018-04-30</td>
<td>7906, 2019-04-23</td>
</tr>
<tr>
<td></td>
<td>NRTI</td>
<td>01.02.00</td>
<td>5336, 2018-10-24</td>
<td>5929, 2018-12-05</td>
</tr>
<tr>
<td>L2_AER_AI</td>
<td></td>
<td>01.02.02</td>
<td>5932, 2018-12-05</td>
<td>7518, 2019-03-27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.00</td>
<td>7519, 2019-03-27</td>
<td>7999, 2019-04-30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.01</td>
<td>8000, 2019-04-30</td>
<td>8906, 2019-07-03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.02</td>
<td>8906, 2019-07-03</td>
<td>16258, 2020-12-02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.04.00</td>
<td>16259, 2020-12-02</td>
<td>19307, 2021-07-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.00</td>
<td>19308, 2021-07-05</td>
<td>21222, 2021-11-17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.01</td>
<td>21223, 2021-11-17</td>
<td>current version</td>
</tr>
<tr>
<td></td>
<td>OFFL</td>
<td>01.02.00</td>
<td>5236, 2018-10-17</td>
<td>5832, 2018-11-28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.02.02</td>
<td>5833, 2018-11-28</td>
<td>7424, 2019-03-20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.00</td>
<td>7425, 2019-03-20</td>
<td>7906, 2019-04-23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.01</td>
<td>7907, 2019-04-23</td>
<td>8814, 2019-06-26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.02</td>
<td>8815, 2019-06-26</td>
<td>16212, 2020-11-29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.04.00</td>
<td>16213, 2020-11-29</td>
<td>19257, 2021-07-01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.00</td>
<td>19258, 2021-07-01</td>
<td>21187, 2021-11-14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.01</td>
<td>21188, 2021-11-14</td>
<td>current version</td>
</tr>
<tr>
<td></td>
<td>NRTI</td>
<td>01.04.00</td>
<td>17155, 2021-01-03</td>
<td>19307, 2021-07-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.00</td>
<td>19308, 2021-07-05</td>
<td>21222, 2021-11-17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.01</td>
<td>21223, 2021-11-17</td>
<td>current version</td>
</tr>
<tr>
<td></td>
<td>L2_AER_LH</td>
<td>01.03.00</td>
<td>7425, 2019-03-20</td>
<td>7906, 2019-04-23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.01</td>
<td>7907, 2019-04-23</td>
<td>8814, 2019-06-26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.03.02</td>
<td>8815, 2019-06-26</td>
<td>16212, 2020-11-29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>01.04.00</td>
<td>16213, 2020-11-29</td>
<td>19257, 2021-07-01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.02.00</td>
<td>19258, 2021-07-01</td>
<td>21187, 2021-11-14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02.03.01</td>
<td>21188, 2021-11-14</td>
<td>current version</td>
</tr>
<tr>
<td></td>
<td>RPRO</td>
<td>01.03.01</td>
<td>2818, 2018-04-30</td>
<td>7424, 2019-03-20</td>
</tr>
</tbody>
</table>

Table 1 – SSP TROPOMI data products and processor versions (NRTI near-real-time and OFFL off-line). Note 1: the operational phase (E2) of the SSP TROPOMI mission starts with orbit #2818. Note 2: RPRO 01.03.01 and 01.03.02 have been used to fill gaps in the 01.02.02 RPRO, therefore processor start and end dates are not sequential.
## Representative Quality Indicators

Representative values of key quality indicators (bias and dispersion vs. reference measurements, and special features) have been derived for the following S5P operational data products on the basis of the validation results reported in this document:

<table>
<thead>
<tr>
<th>Product ID</th>
<th>Stream</th>
<th>Product</th>
<th>Bias</th>
<th>Dispersion</th>
<th>Special features</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2_O3</td>
<td>NRTI</td>
<td>O₃ column</td>
<td>0.8%</td>
<td>2.5%</td>
<td>Larger dispersion over snow/ice due to coarse surface albedo climatology (up to but excluding v02.01.xx, which has a dynamic determination of surface albedo). Potentially some increase in overall bias (+0.7%) since v02.02.01 (5 July 2021).</td>
</tr>
<tr>
<td></td>
<td>OFFL</td>
<td>O₃ column</td>
<td>0.3%</td>
<td>2%</td>
<td>Some increase in overall bias (+0.5 to +1.5%) since v02.02.01 (1 July 2021).</td>
</tr>
<tr>
<td>L2_O3_TCL</td>
<td>OFFL</td>
<td>O₃ tropospheric column</td>
<td>+17%</td>
<td>26%</td>
<td>Geographical imprints of sampling-related biases. Seasonal change of the bias. More elevated positive bias in Atlantic region during biomass burning season.</td>
</tr>
<tr>
<td>L2_O3_PR</td>
<td>NRTI</td>
<td>O₃ profile</td>
<td>10%</td>
<td>5-20%</td>
<td>Bias below 10 % in the troposphere up to UTLS, and higher dispersion. Vertically oscillating bias of 5-10 % (positive to negative) in the stratosphere, with a smaller dispersion.</td>
</tr>
<tr>
<td></td>
<td>OFFL</td>
<td>O₃ profile</td>
<td>10%</td>
<td>5-20%</td>
<td></td>
</tr>
<tr>
<td>L2_NO2</td>
<td>NRTI</td>
<td>NO₂ troposphere</td>
<td>-37%</td>
<td>2.6 Pmolec/cm²</td>
<td>The bias and dispersion sorted by column amount: Troposphere ([-2 \text{ Pmolec/cm}^2]) +18% ((0.7 \text{ Pmolec/cm}^2)), ([-15 \text{ Pmolec/cm}^2]) -46% ((7.3 \text{ Pmolec/cm}^2)). Total ([+/- 6 \text{ Pmolec/cm}^2]): +4% ((1.0 \text{ Pmolec/cm}^2)) and -21% ((2.2 \text{ Pmolec/cm}^2)). The products improve with later versions. VDAF-AVS tropospheric bias decreases to -16% for PAL reprocessed data.</td>
</tr>
<tr>
<td></td>
<td>OFFL</td>
<td>NO₂ stratosphere</td>
<td>-34%</td>
<td>2.6 Pmolec/cm²</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RPRO</td>
<td>NO₂ stratosphere</td>
<td>-5%</td>
<td>0.3 Pmolec/cm²</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RPRO</td>
<td>NO₂ total</td>
<td>-7%</td>
<td>1.5 Pmolec/cm²</td>
<td></td>
</tr>
<tr>
<td>L2_HCHO</td>
<td>NRTI</td>
<td>HCHO, low</td>
<td>+28%</td>
<td>9 Pmolec/cm²</td>
<td>Bias and dispersion depend on column amount: ([&lt;2.5 \text{ Pmolec/cm}^2]) positive bias, low dispersion, ([&gt;8 \text{ Pmolec/cm}^2]) negative bias, high dispersion.</td>
</tr>
<tr>
<td></td>
<td>OFFL</td>
<td>HCHO, high</td>
<td>-29%</td>
<td>25 Pmolec/cm²</td>
<td></td>
</tr>
<tr>
<td>L2_SO2</td>
<td>NRTI</td>
<td>SO₂ column</td>
<td>0.2 DU</td>
<td>0.2 DU</td>
<td>Lack of validation stations in areas with high SO₂.</td>
</tr>
<tr>
<td></td>
<td>OFFL</td>
<td>SO₂ column</td>
<td>0.2 DU</td>
<td>0.2 DU</td>
<td></td>
</tr>
<tr>
<td>L2_CO</td>
<td>NRTI</td>
<td>CO column</td>
<td>6.5%</td>
<td>5%</td>
<td>Along orbit stripes. High pollution underestimated. 5% SZA dependence of bias. Outliers in SAA and other sporadic locations not filtered by qa_value. Since July 2019 NRTI similar as OFFL. Processor update on July 1, 2021 introduces a de-striped product and a change in spectroscopic parameters Preliminary results indicate that the bias reduces to 2.9%.</td>
</tr>
<tr>
<td></td>
<td>OFFL</td>
<td>CO column</td>
<td>6.5%</td>
<td>5%</td>
<td></td>
</tr>
</tbody>
</table>

**ATM MPC**
### Table 2 – Representative quality indicators (bias, dispersion and special features) derived from the validation of the SSP TROPOMI data products listed in the Table 1, valid for all processor versions unless stated differently.

<table>
<thead>
<tr>
<th>Product ID</th>
<th>Stream</th>
<th>Product</th>
<th>Bias</th>
<th>Dispersion</th>
<th>Special features</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2_CH4</td>
<td>OFFL</td>
<td>CH4 column</td>
<td>-0.03%</td>
<td>0.63%</td>
<td>Along orbit stripes. Underestimation at low albedo. Remaining outliers with qa_value &gt;0.5. 1-4% seasonal and SZA dependence of bias. Lower amount of pixels with qa_value &gt;0.5 between March 11 2020 and July 1 2021 due to changed cloud data. Processor update on November 14 2021 produces data over the ocean (sun-glint), updated spectroscopy, a-posteriori bias correction independent of any reference data.</td>
</tr>
<tr>
<td>L2_CLOUD</td>
<td>NRTI</td>
<td>CALv1 CTH (high)</td>
<td>-30%</td>
<td>2 km</td>
<td>Low clouds: CLOUDNET CTH&lt;4km; high clouds: CLOUDNET CTH&gt;4km. Snow/ice albedo degrades retrievals, improved with version 02.01.03. Occurrence of C(T)H equal to surface height at low cloud fraction, improved with version upgrade. Across track CTH and CF pattern and North-South cloud albedo pattern, improved with version 02.01.03. Cloud fraction: lower dispersion between CLOUD 02.02 and FRESCO 02.02.</td>
</tr>
<tr>
<td></td>
<td>OFFL</td>
<td>CALv2 CTH (high)</td>
<td>-40%</td>
<td>2 km</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CALv1 CTH (low)</td>
<td>-15%</td>
<td>0.8 km</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CALv2 CTH (low)</td>
<td>-15%</td>
<td>0.6 km</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CRBv1 CH (high)</td>
<td>-20%</td>
<td>1.5 km</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CRBv2 CH (high)</td>
<td>-25%</td>
<td>1.8 km</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CRBv1 CH (low)</td>
<td>-35%</td>
<td>0.6 km</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CRBv2 CH (low)</td>
<td>-40%</td>
<td>0.5 km</td>
<td></td>
</tr>
<tr>
<td>L2_AER_AI</td>
<td>NRTI</td>
<td>aerosol index</td>
<td>-1.1 AI unit</td>
<td>0.1 AI unit</td>
<td>Negative bias exceeding 1 AI unit after March 2019, attributed to irradiance data degradation. The issue was resolved with the L1B processor upgrade to version 2 in July 2021.</td>
</tr>
<tr>
<td></td>
<td>OFFL</td>
<td>aerosol index</td>
<td>-1.1 AI unit</td>
<td>0.1 AI unit</td>
<td></td>
</tr>
<tr>
<td>L2_AER_LH</td>
<td>OFFL</td>
<td>aerosol layer height</td>
<td>50 hPa</td>
<td>100 hPa</td>
<td>Over ocean only. Larger bias and dispersion expected over land.</td>
</tr>
</tbody>
</table>
Table of Contents

Document Information .................................................................................................................. 4
Executive Summary ..................................................................................................................... 5
Processing Baseline Identification ............................................................................................... 12
Representative Quality Indicators ............................................................................................... 15
Table of Contents .......................................................................................................................... 17
1 Introduction .................................................................................................................................... 23
  1.1 Background information on Sentinel-5 Precursor TROPOMI ............................................. 23
  1.2 Atmospheric Mission Performance Cluster – Routine Operations Validation Service ............. 23
  1.3 Purpose, scope and outline of this document ..................................................................... 24
2 S5P Data Quality Requirements ............................................................................................... 25
3 Validation Results: L1B_RA and L1B_IR ............................................................................... 26
  3.1 L1B products ......................................................................................................................... 26
  3.2 Recommendations for data usage followed ....................................................................... 26
  3.3 Validation approach ............................................................................................................. 27
  3.4 Validation of L1B NRTI ....................................................................................................... 27
  3.5 Validation of L1B OFFL ........................................................................................................ 27
4 Validation Results: L2_O3 ......................................................................................................... 28
  4.1 L2_O3 products and requirements .................................................................................... 28
  4.2 Validation approach ............................................................................................................. 28
    4.2.1 Ground-based networks ............................................................................................... 28
    4.2.2 Satellites ..................................................................................................................... 28
    4.2.3 Field campaigns and modelling support ..................................................................... 28
  4.3 Validation of L2_O3 NRTI ................................................................................................. 28
    4.3.1 Recommendations for data usage followed ................................................................. 28
    4.3.2 Status of validation ...................................................................................................... 29
    4.3.3 Bias .............................................................................................................................. 30
      4.3.3.1 Switch to smaller ground pixel size and to UPAS processor v2.1.3 and v2.2.1 .................. 31
    4.3.4 Dispersion .................................................................................................................... 32
    4.3.5 Dependence on influence quantities .......................................................................... 33
    4.3.6 Short term variability ................................................................................................... 34
    4.3.7 Geographical patterns ................................................................................................. 35
    4.3.8 Other features .............................................................................................................. 35
  4.4 Validation of L2_O3 OFFL ................................................................................................. 36
    4.4.1 Recommendations for data usage followed ................................................................. 36
    4.4.2 Status of validation ...................................................................................................... 36
    4.4.3 Bias .............................................................................................................................. 37
      4.4.3.1 Switch to smaller ground pixel size and to UPAS processor v2.1.3 and v2.2.1 .................. 38
    4.4.4 Dispersion .................................................................................................................... 39
    4.4.5 Dependence on influence quantities .......................................................................... 40
    4.4.6 Short term variability ................................................................................................... 41
5 Validation Results: L2_O3_TCL........................................................................ 42
  5.1 L2_O3_TCL products and requirements ............................................... 42
  5.2 Validation approach .............................................................................. 42
    5.2.1 Ground-based networks ................................................................ 42
    5.2.2 Satellites ......................................................................................... 42
    5.2.3 Field campaigns and modelling support ......................................... 42
  5.3 Validation of L2_O3_TCL OFFL (CCD) ......................................................... 43
    5.3.1 Recommendations for data usage followed .................................... 43
    5.3.2 Status of validation ......................................................................... 43
    5.3.3 Bias ............................................................................................... 44
    5.3.4 Dispersion ...................................................................................... 44
    5.3.5 Dependence on influence quantities .................................................. 44
    5.3.6 Seasonal cycle and shorter term variability ....................................... 45
    5.3.7 Geographical patterns .................................................................... 45
    5.3.8 Preliminary validation of UPAS processor upgrade to version 02.02.01 45
    5.3.9 Other features ................................................................................ 46

6 Validation Results: L2_O3_PR .................................................................... 50
  6.1 L2_O3_PR products and requirements .................................................... 50
  6.2 Validation approach .............................................................................. 50
    6.2.1 Ground-based networks ................................................................ 50
      6.2.1.1 Balloon-borne ozonesonde ....................................................... 50
      6.2.1.2 Differential absorption ozone lidars ........................................ 51
    6.2.2 Satellite intercomparisons ................................................................. 51
    6.2.3 Analysis of information content ...................................................... 51
    6.2.4 Analysis of daily global maps ............................................................ 51
    6.2.5 Parameter correlation checks .......................................................... 51
    6.2.6 Field campaigns and modelling support ......................................... 52
  6.3 Validation of L2_O3_PR NRTI ................................................................. 52
    6.3.1 Recommendations for data usage followed .................................... 52
    6.3.2 Status of validation ......................................................................... 52
    6.3.3 Vertical sensitivity, resolution and registration .................................. 54
    6.3.4 Bias ............................................................................................... 54
    6.3.5 Dispersion ...................................................................................... 54
    6.3.6 Chi-square tests ............................................................................ 54
    6.3.7 Dependence on influence quantities ................................................. 55
    6.3.8 Short term variability ...................................................................... 55
    6.3.9 Geographical patterns .................................................................... 55
    6.3.10 Other features ................................................................................ 55
  6.4 Validation of L2_O3_PR OFFL ................................................................. 57
    6.4.1 Recommendations for data usage followed .................................... 57
6.4.2 Status of validation ................................................................. 57
6.4.3 Vertical sensitivity, resolution and registration.......................... 58
6.4.4 Bias .................................................................................. 58
6.4.5 Dispersion ........................................................................... 59
6.4.6 Chi-square tests .................................................................... 59
6.4.7 Dependence on influence quantities ........................................ 59
6.4.8 Short term variability ............................................................ 59
6.4.9 Geographical patterns ............................................................ 60
6.4.10 Other features ..................................................................... 60

7 Validation Results: L2_NO2 .............................................................. 62
7.1 L2_NO2 products and requirements ............................................. 62
7.2 Validation approach ................................................................... 62
7.2.1 Ground-based monitoring networks ......................................... 62
7.2.2 Satellites ............................................................................ 64
7.2.3 Field campaigns and modelling support .................................... 64
7.3 Validation of L2_NO2 ................................................................. 65
7.3.1 Recommendations for data usage ............................................. 65
7.3.2 Status of validation ............................................................... 65
7.3.3 Tropospheric NO2 column ...................................................... 66
7.3.4 Stratospheric NO2 column ...................................................... 74
7.3.5 Total NO2 column ................................................................. 79
7.4 Equivalence of L2_NO2 NRTI and OFFL products ....................... 83
7.4.1 Tropospheric NO2 Column NRTI vs OFFL ............................... 83
7.4.2 Stratospheric NO2 Column NRTI vs OFFL ............................... 84
7.4.3 Total NO2 Column NRTI vs OFFL ........................................... 85
7.5 Internal consistency of the NO2 validation results ......................... 86
7.5.1 NO2 absorption cross-sections ............................................... 86
7.5.2 ZSL-DOAS and PGN with low pollution level (mountain-top, arctic) .......................................................... 86
7.5.3 Stations with multiple instruments (different geometries) ............ 87
7.5.4 Consistency between MAX-DOAS and PGN network results .......... 87

8 Validation Results: L2_HCHO .......................................................... 89
8.1 L2_HCHO products and requirements ........................................ 89
8.2 Validation approach ................................................................. 89
8.2.1 Ground-based monitoring networks ....................................... 89
8.2.2 Satellites ............................................................................ 91
8.2.3 Field campaigns and modelling support .................................... 91
8.3 Validation of L2_HCHO ............................................................. 91
8.3.1 Recommendations for data usage ............................................ 91
8.3.2 Status of validation ............................................................... 91
8.3.3 Bias .................................................................................. 92
8.3.4 Dispersion ........................................................................... 98
8.3.5 Dependence on influence quantities ....................................... 100
8.3.2 Seasonal and short term variability ................................................................. 100
8.3.7 Geographical patterns .................................................................................... 101
8.3.8 Impact of UPAS processor upgrades up to version 02.03.00 ............................... 102
8.4 Equivalence of L2_HCHO NRTI and OFFL products ......................................... 103
8.4.1 Bias ................................................................................................................. 104
8.4.2 Dispersion ...................................................................................................... 105

9 Validation Results: L2_SO2 .......................................................... 106
9.1 L2_SO2 products and requirements .................................................................. 106
9.2 Validation approach .......................................................................................... 106
9.2.1 Ground-based networks ................................................................................ 106
9.2.2 Satellites ...................................................................................................... 106
9.2.3 Field campaigns and modelling support ....................................................... 106
9.2.4 Test of the expectation of zero SO₂ SCDs (within detection limit) outside volcanic plumes and strongly polluted regions ............................................. 107
9.3 Validation of L2_SO2 NRTI ............................................................................. 107
9.3.1 Recommendations for data usage followed ................................................ 107
9.3.2 Status of validation ....................................................................................... 107
9.3.3 Bias ................................................................................................................. 111
9.3.4 Dispersion .................................................................................................... 111
9.3.5 Dependence on influence quantities ............................................................. 113
9.3.6 Short term variability .................................................................................... 113
9.3.7 Geographical patterns .................................................................................. 113
9.3.8 Other features .............................................................................................. 113
9.4 Equivalence of L2_SO2 NRTI and OFFL products .......................................... 113

10 Validation Results: L2_CO ............................................................. 114
10.1 L2_CO products and requirements ............................................................... 114
10.2 Validation approach ....................................................................................... 114
10.2.1 Ground-based networks ............................................................................. 114
10.2.2 Satellites .................................................................................................... 114
10.2.3 Field campaigns and modelling support ................................................... 115
10.3 Validation of L2_CO OFFL ................................................................. 115
10.3.1 Recommendations for data usage followed ............................................. 115
10.3.2 Status of validation .................................................................................... 115
10.3.3 Bias .............................................................................................................. 116
10.3.4 Dispersion .................................................................................................. 120
10.3.5 Dependence on influence quantities ......................................................... 120
10.3.6 Short term variability .................................................................................. 122
10.3.7 Geographical patterns ............................................................................... 124
10.3.8 Other features ........................................................................................... 125
10.4 Equivalence of L2_CO OFFL and NRTI products ........................................... 126
10.5 Accuracy estimate for processor 02.02 ......................................................... 126
10.6 Comparison of L2_CO OFFL standard and destriped products .................... 127
11 Validation Results: L2_CH4 ................................................................. 128
  11.1 L2_CH4 products and requirements ........................................... 128
  11.2 Validation approach .................................................................. 128
    11.2.1 Ground-based networks ...................................................... 128
    11.2.2 Satellites ........................................................................... 128
    11.2.3 Other TROPOMI XCH4 products ......................................... 129
    11.2.4 Field campaigns and modelling support ................................ 130
  11.3 Validation of L2_CH4 OFFL ........................................................ 130
    11.3.1 Recommendations for data usage followed .......................... 130
    11.3.2 Status of validation ............................................................ 131
    11.3.3 Bias .................................................................................. 132
    11.3.4 Dispersion ......................................................................... 136
    11.3.5 Dependence on influence quantities .................................... 136
    11.3.6 Short term variability .......................................................... 137
    11.3.7 Geographical patterns ......................................................... 138
    11.3.8 Change in Cloud Mask ........................................................ 139
    11.3.9 Other features .................................................................... 140
  11.4 First assessment of L2_CH4 OFFL processor update to version 02.02.00 ................................................................. 142
  11.5 Validation of L2_CH4 OFFL sun-glint retrievals ............................ 142
12 Validation Results: L2_CLOUD .......................................................... 145
  12.1 L2_CLOUD products and requirements ....................................... 145
  12.2 Validation approach .................................................................. 145
    12.2.1 Ground-based networks ...................................................... 145
    12.2.2 Satellites ........................................................................... 146
    12.2.3 Alternative S5P cloud algorithms ....................................... 147
    12.2.4 Field campaigns and modelling support ............................... 147
  12.3 Validation of L2_CLOUD OFFL .................................................... 148
    12.3.1 Recommendations for data usage followed .......................... 148
    12.3.2 Status of validation ............................................................ 148
    12.3.3 Radiometric cloud fraction (L2_CLOUD CAL & L2_CLOUD CRB) ................................................................. 148
    12.3.4 Cloud top height (L2_CLOUD CAL) and cloud height (L2_CLOUD CRB) ......................................................... 153
  12.4 Comparison of L2_CLOUD NRTI and OFFL products .................. 165
13 Validation Results: L2_AER_AI .......................................................... 166
  13.1 L2_AER_AI products and requirements ....................................... 166
  13.2 Validation approach .................................................................. 166
    13.2.1 Ground-based networks ...................................................... 166
    13.2.2 Satellites ........................................................................... 166
    13.2.3 Field campaigns and modelling support ............................... 166
  13.3 Validation of L2_AER_AI NRTI ..................................................... 167
    13.3.1 Recommendations for data usage followed .......................... 167
    13.3.2 Status of validation ............................................................ 167
    13.3.3 Bias .................................................................................. 171
13.3.4 Dispersion ......................................................................................................... 171
13.3.5 Dependence on influence quantities ............................................................... 171
13.3.6 Short term variability ..................................................................................... 171
13.3.7 Geographical patterns ................................................................................... 175
13.3.8 Other features ............................................................................................... 175
13.4 Equivalence of L2_AER_AI NRTI and OFFL products ....................................... 175
14 Validation Results: L2_AER_LH ......................................................................... 176
14.1 L2_AER_LH products and requirements ............................................................. 176
14.2 Validation approach ......................................................................................... 176
14.2.1 Ground-based networks .............................................................................. 176
14.2.2 Satellites ...................................................................................................... 177
14.2.3 Field campaigns and modelling support ....................................................... 177
14.3 Validation of L2_AER_LH ................................................................................. 177
14.3.1 Recommendations for data usage followed ................................................ 177
14.3.2 Status of validation ...................................................................................... 178
14.3.3 Validation approach .................................................................................... 185
14.3.4 Case studies ................................................................................................ 187
14.3.5 Bias .............................................................................................................. 189
14.3.6 Dispersion .................................................................................................... 189
14.3.7 Dependence on influence quantities ............................................................ 190
14.3.8 Short term variability .................................................................................. 193
14.3.9 Geographical patterns ................................................................................. 193
14.3.10 Other features ............................................................................................ 194
15 References ........................................................................................................... 197
15.1 Reference documents ...................................................................................... 197
15.2 Peer-reviewed articles ..................................................................................... 197
15.3 Electronic references ....................................................................................... 203
16 Acknowledgements ............................................................................................. 205
16.1 ATM-MPC SSP Routine Operations Validation Service ......................................... 205
16.2 SSP validation facilities .................................................................................. 206
16.3 Validation data ................................................................................................ 206
16.4 Agency support ................................................................................................ 208
17 Terms, definitions and abbreviated terms ............................................................. 209
17.1 Terms and definitions ...................................................................................... 209
17.2 Acronyms and abbreviations ........................................................................... 210
1 Introduction

1.1 Background information on Sentinel-5 Precursor TROPOMI

TROPOnospheric Monitoring Instrument (TROPOMI) [ER_TROPOMI] is the unique payload of the ESA/Copernicus Sentinel-5 Precursor mission (S5P) launched on October 13, 2017. The prime function of TROPOMI is to monitor the global distribution of atmospheric trace gases and aerosols for a better understanding of air quality, the ozone layer, atmospheric chemistry and transport, ultraviolet radiation, and climate change. The instrument is a nadir-viewing hyperspectral spectrometer measuring, in the ultraviolet-visible (270-495 nm), near infrared (675-775 nm) and shortwave infrared (2305-2385 nm), the solar radiation scattered by the Earth’s atmosphere and reflected by the Earth’s surface and by clouds, as well as solar spectral irradiance. Daily coverage at the high horizontal resolution of 7 x 3.5 km² before and 5.5 x 3.5 km² after the operations switch to smaller ground pixel size activated on the 6th of August 2019, is accomplished thanks to a Sun-synchronous polar orbit (equator crossing time of 13:30 local solar time) and a wide swath width of 2600 km across track. From the TROPOMI radiometric measurements of Earth’s radiance and solar irradiance, on-ground data processors retrieve the atmospheric abundance of ozone (O₃), nitrogen dioxide (NO₂), formaldehyde (HCHO), sulphur dioxide (SO₂), carbon monoxide (CO), methane (CH₄), as well as cloud and aerosol properties.

The S5P mission is a key component of the space segment of the European Earth Observation programme Copernicus [ER_Copernicus]. As such, it has an operational and service-oriented vocation. With a 7-year nominal operation lifetime, the S5P mission aims at filling in the anticipated observational gap of key atmospheric composition data between, from one part, Envisat SCIAMACHY (operational in 2002-2012), EOS-Aura OMI (operational since 2004) and the EUMETSAT EPS MetOp GOME-2 series (initiated in 2006, with the latest MetOp-C launched in November 2018), and from the other part, the upcoming series of Copernicus Sentinel-4 and Sentinel-5 missions scheduled for 2024-2045.

1.2 Atmospheric Mission Performance Cluster – Routine Operations Validation Service

Procured by an international consortium contracted by the European Space Agency (ESA), the Copernicus Atmospheric Mission Performance Cluster (ATM-MPC) provides an operational service-based response to the S5P mission requirements for quality control, calibration, validation and end-to-end system performance monitoring during the Routine Operations phase of the S5P mission.

In-flight calibration and characterisation of the TROPOMI instrument, long-term monitoring of the instrument sensor performance and ageing, and routine Quality Control (QC) of the operational Level-1 (radiometric) and Level-2 (geophysical) data products are coordinated by the Royal Dutch Meteorological Institute (KNMI), and documented on the TROPOMI Portal for Instrument and Calibration [ER_MPS] and the TROPOMI Portal for Level-2 Quality Control [ER_L2QC].

Geophysical validation of the operational Level-1 and Level-2 data products is coordinated by the Royal Belgian Institute for Space Aeronomy (BIRA-IASB), and documented on the Portal of the TROPOMI Validation Data Analysis Facility (VDAF) [ER_VDAF]. The TROPOMI routine operations validation service makes use of Fiducial Reference Measurements (FRM) and other correlative data of documented quality (ground-based and satellite measurements, dedicated field campaigns), to assess the overall quality, the compliance with mission requirements and the validity of uncertainty estimates of the TROPOMI data products. This service monitors validation results on a cyclic basis and updates every three months the present Routine Operations Consolidated Validation Report (ROCR). It also contributes quality assessment support to the continuous evolution of the data processors.
1.3 Purpose, scope and outline of this document

The present document (DI-MPC-ROCVR) reports consolidated validation results for the S5P TROPOMI Level-1 and Level-2 operational data products. This report has been produced by the ATM-MPC Routine Operations Validation Service. It integrates validation results from the MPC Validation Data Analysis Facility (VDAF) consortium (Table 23) with support from other activities and dedicated field campaigns documented on the TROPOMI website [ER_TROPOMI], as well as ad hoc contributions from SSP Validation Team (SSPVT) AO projects [ER_SSPVT].

Updated with a trimestral frequency, SSP data quality information provided in this document supersedes that provided in previous versions. It complements SSP data quality information provided in the Product Readme Files (PRFs) attached to SSP data products released publicly. For details and for recommendations for data usage, data users are encouraged to read the PRF, Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) associated with the data products, all available on the Copernicus Sentinel Portal for SSP products and algorithms [ER_CoperATBD] and also on the TROPOMI Portal [ER_TROPOMI].

This ROCVR update #15 reports quality information for the S5P operational data products acquired from April 2018 until May 2022. It includes validation results for version 02.02.00 and 02.03.01 of the NL-L2 processor suite (NO₂, CO, CH₄, AAI, and ALH) and version 02.02.01 of the UPAS processor suite (total and tropospheric O₃ columns, HCHO, SO₂, and CLOUD data). Validation of the ozone profile data (NL-L2 processor v02.03.01) is also consolidated. This document is structured as follows:

| Document Information | .......................................................... | .......................................................... | .......................................................... | .......................................................... | .......................................................... | .......................................................... | .......................................................... | .......................................................... | .......................................................... | .......................................................... | .......................................................... | .......................................................... |
|----------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| Executive Summary    | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| Processing Baseline Identification | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| Representative Quality Indicators | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| Table of Contents   | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 1 Introduction       | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 2 SS5 Data Quality Requirements | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 3 Validation Results: L1B_RA and L1B_IR | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 4 Validation Results: L2_O3............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 5 Validation Results: L2_O3_TCL............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 6 Validation Results: L2_O3_PR............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 7 Validation Results: L2_NO2............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 8 Validation Results: L2_HCHO............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 9 Validation Results: L2_SO2............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 10 Validation Results: L2_CO............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 11 Validation Results: L2_CH4............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 12 Validation Results: L2_CLOUD............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 13 Validation Results: L2_AER_AI............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 14 Validation Results: L2_AER_LH............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 15 References ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 16 Acknowledgements ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |
| 17 Terms, definitions and abbreviated terms ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. | ............................................................................................................................................. |

ATM MPC
2 S5P Data Quality Requirements

Validation results can be interpreted to evaluate whether or not S5P Level 2 data products meet user requirements. Targets for key quality indicators of the S5P Level 2 data products have been formulated in the S5P Geophysical Validation Requirements document ([S5PVT-Req], Page 19) and the S5P Cal/Val Plan for the Operational Phase ([S5P-CSCOP], Page 14). Evolution of these requirements is supported by the Sentinel-5p Quality Working Group (QWG), who agreed (i) to adopt for tropospheric ozone column data the requirements expressed by the Climate Research Group (CRG) within ESA’s Ozone_cci project, (ii) to revise requirements for SO\(_2\) column data, and (iii) to provide maximum values of the estimates instead of ranges. Refined requirements were adopted for three different cases of SO\(_2\) column regimes. Expressed in terms of measurement bias (estimate of the systematic measurement error) and uncertainty (measurement uncertainty, that is, dispersion of the quantity values being attributed to the measurand), these targets are reproduced hereafter in Table 3. Quality targets are typical of several known applications; nevertheless, it always remains the uttermost responsibility of any users to check the fitness of the S5P data for their own purpose, with respect to their own particular requirements.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>L2_O3</td>
<td>Total O(_3)</td>
<td>total column</td>
<td>5%</td>
<td>2.5%</td>
</tr>
<tr>
<td>L2_O3_PR</td>
<td>O(_3) profile (incl. troposphere)</td>
<td>6 km</td>
<td>30%</td>
<td>10%</td>
</tr>
<tr>
<td>L2_O3_TCL</td>
<td>O(_3) tropospheric column</td>
<td>tropospheric column</td>
<td>25%</td>
<td>25%</td>
</tr>
<tr>
<td>L2_NO2</td>
<td>NO(_2) tropospheric column</td>
<td>tropospheric column</td>
<td>50%</td>
<td>0.7 Pmolec.cm(^2)</td>
</tr>
<tr>
<td>L2_NO2</td>
<td>NO(_2) stratospheric column</td>
<td>stratospheric column</td>
<td>10%</td>
<td>0.5 Pmolec.cm(^2)</td>
</tr>
<tr>
<td>L2_SO2</td>
<td>Total SO(_2)</td>
<td>total column</td>
<td>0.5 DU</td>
<td>1 DU</td>
</tr>
<tr>
<td>L2_SO2</td>
<td>Enhanced SO(_2) (SCD &lt;1.5 DU)</td>
<td>total column</td>
<td>0.5 DU</td>
<td>1 DU</td>
</tr>
<tr>
<td>L2_SO2</td>
<td>Enhanced SO(_2) (SCD &gt;1.5 DU)</td>
<td>total column</td>
<td>30%</td>
<td>30%</td>
</tr>
<tr>
<td>L2_HCHO</td>
<td>Total HCHO</td>
<td>total column</td>
<td>80%</td>
<td>12 Pmolec.cm(^2)</td>
</tr>
<tr>
<td>L2_CO</td>
<td>Total CO</td>
<td>total column</td>
<td>15%</td>
<td>10%</td>
</tr>
<tr>
<td>L2_CH4</td>
<td>Total CH(_4)</td>
<td>total column</td>
<td>1.5%</td>
<td>1%</td>
</tr>
<tr>
<td>L2_CLOUD</td>
<td>Cloud Fraction</td>
<td>total column</td>
<td>20%</td>
<td>0.05</td>
</tr>
<tr>
<td>L2_CLOUD</td>
<td>Cloud Height (pressure)</td>
<td>total column</td>
<td>20%</td>
<td>0.5km (P&lt;30hPa)</td>
</tr>
<tr>
<td>L2_CLOUD</td>
<td>Cloud albedo (optical thickness)</td>
<td>total column</td>
<td>20%</td>
<td>0.05 (10)</td>
</tr>
<tr>
<td>L2_AER_AI</td>
<td>Aerosol Absorbing Index</td>
<td>total column</td>
<td>1 AAI</td>
<td>0.1 AAI</td>
</tr>
<tr>
<td>L2_AER_ALH</td>
<td>Aerosol Layer Height</td>
<td>total column</td>
<td>100 hPa</td>
<td>50 hPa</td>
</tr>
</tbody>
</table>

Table 3 – Data quality targets for the operational Sentinel-5 Precursor TROPOMI Level 2 data products: vertical resolution, and measurement uncertainty components associated with systematic and random effects, respectively (adapted by Sentinel-5p QWG from [S5PVT-Req] and [S5P-CSCOP]). SCD: slant column density, before the conversion to vertical column density using an air mass factor.
3 Validation Results: L1B_RA and L1B_IR

3.1 L1B products

This Section reports on the validation of the S5P TROPOMI L1B products identified in Table 4, including the implementation of the L1B processor update in July 2021. Current conclusions are based on the limited amount of version 2 data available at the time of this first analysis, and on a period covered by this dataset. The conclusions summarized hereafter need to be confirmed by a larger amount of colocations, and extended over at least a full year of data, hence, a full cycle of key influence quantities, in order to enable detection and quantification of potential patterns, dependences, seasonal cycles and longer-term features.

Table 4: Identification of the S5P TROPOMI L1B products evaluated in this Section.

<table>
<thead>
<tr>
<th>Product</th>
<th>Stream</th>
<th>Version</th>
<th>In operation from</th>
<th>In operation until</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1B_RA1/.../8</td>
<td>01.00.00</td>
<td>orbit 2818, 2018-04-30</td>
<td>19257, 2021-07-01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>02.00.00</td>
<td>19258, 2021-07-01</td>
<td>current version</td>
<td></td>
</tr>
<tr>
<td>L1B_IR_UVN/SIR</td>
<td>01.00.00</td>
<td>orbit 2818, 2018-04-30</td>
<td>19257, 2021-07-01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>02.00.00</td>
<td>19258, 2021-07-01</td>
<td>current version</td>
<td></td>
</tr>
</tbody>
</table>

Note: The operational phase (E2) of the S5P TROPOMI mission starts with orbit #02818.

3.2 Recommendations for data usage followed

An overview of the Sentinel-5p mission, the TROPOMI instrument and the algorithms for producing the L1b data products can be found in the Algorithm Theoretical Basis Document [ATBD]. Details of the data format are provided in the Input/Output Data Specification [IODS]. The metadata contained in the L1b data products are described in the Metadata Specification [MDS]. All these documents are available on https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms.

For Level 2 processing and related validation, the following additional notices apply:

- The L0-1b data processor annotates the data with quality assessment data in the fields spectral_channel_quality, measurement_quality and ground_pixel_quality. Level 2 developers are strongly encouraged to observe these quality fields in their retrievals and exclude flagged data as needed.

- All eight bands are processed individually in the L0-1b data processor. In case of missing data, for example in case of data dropouts during downlinks, this does not necessarily impact all bands (to the same extent). This means that a scanline can be missing for some bands, where it is not missing for other bands. When combining data from multiple bands, Level 2 algorithms should therefore always check and match the delta_time for these data and, in case of non-co-registered bands, the geolocation as well.

- For calculating reflectance from the radiance products, it is recommended to use the irradiance product with the sensing time close to the sensing time of the radiance product.
3.3 Validation approach

In-flight calibration and characterisation of the TROPOMI instrument, long-term monitoring of the instrument sensor performance and ageing, and routine Quality Control (QC) of the operational L1B data products are reported continuously on the TROPOMI Portal for Instrument and Calibration [ER_MPS].

The SSP/TROPOMI Level 1b products have been compared to model outputs, to other satellite measurements, and to ground based measurements.

3.4 Validation of L1B NRTI

The near-real time L1b products are not distributed to users, and they are not validated separately. NRTI products use the same L01b data processor algorithms, and can only differ when the Calibration Key Data (CKD) used differs from OFFL. Currently no CKD is dynamically updated in OFFL, and hence no difference exists between NRTI and OFFL.

3.5 Validation of L1B OFFL

The validation of the wavelength assignment of the L1B_UVN version 2 products shows agreement of within 0.01 nm, which is within the pre-launch calibration accuracy. The reflectance in bands 1-3 is 1 % to 3 % lower than TOMS and the used ice radiance model. The radiance in bands 1-3 is up to 5 % smaller than OMPS radiance, above 320 nm this is a wavelength independent bias. Below 320 nm, the wavelength dependence seems to vary with the latitude. In band 1 around 280 nm the radiance deviates more than 10% from OMPS values. The absolute radiometric calibration for UV radiance lacks accuracy and as a result may be updated in the future. In the spectrally overlapping regions of bands 2 and 3 there is a discrepancy of about 2 % in the L1b radiance signals. The radiance in band 6 was compared to model spectra in the continuum around the O2A band. The signal of TROPOMI is 1-2% lower than the model. For bands 1 to 6 (UV, UVIS and NIR) degradation has been observed for the radiance. The degradation is the largest at short wavelengths. The decrease in radiance signal per 1000 orbits is between 0.31% in band 1 and 0.02% in band 6. The degradation is planned to be corrected in a future update of the calibration key data. The absolute and relative radiance radiometry of the SWIR bands were validated using reference stations in Railroad Valley and in the Saharan desert. Current validation results give upper limits of <5% for the absolute calibration and <0.8% for the relative calibration. The absolute irradiance calibration of TROPOMI has been compared to other published solar reference datasets. After an update to the calibration based on OMPS data, the UV and UVIS spectrometers agree within 0-5 % with the references. For extreme swath angles, the deviations are larger in the UV. For the NIR spectrometer the irradiance spectrum is approximately 1.5 %–3.5 % lower than the reference spectra. The SWIR spectrum is approximately 0.6 % lower than the closest reference spectrum.
4 Validation Results: L2_O3

4.1 L2_O3 products and requirements

This Section reports on the validation of the S5P TROPOMI L2_O3 product identified in Table 1. Validation results are discussed with respect to the product quality targets outlined in Table 3. The NRTI and OFFL processors using different retrieval approaches, their respective validation is reported in separate subsections. Special attention is paid to the UPAS processor upgrade to version 02.01.03 – on 16 July 2020 for NRTI and on 13 July 2020 for OFFL – as this included a significant improvement of the surface albedo determination in the NRTI processor (now retrieved instead of taken from a climatology). Also included are the validation results for data since the upgrade of the L1b processor (L2_O3 processor v2.2.1 and above since July 2021).

4.2 Validation approach

4.2.1 Ground-based networks

S5P TROPOMI L2_O3 total ozone column data are routinely compared to reference measurements acquired by instruments contributing to WMO’s Global Atmosphere Watch (GAW): (1) Brewer (Kerr et al., 1981,1988) and (2) Dobson (Basher, 1982) UV spectrophotometers, and (3) NDACC Zenith Scattered Light (ZSL) DOAS UV-Visible spectrometers (Pommereau and Goutail, 1988, Hendrick et al., 2011). Co-locations between S5P TROPOMI and direct-sun (DS) measurements are defined as “pixel contains station”, with a maximum time difference of 3 hours. Note that direct-sun measurements obtained through the NDACC and WOUDC data archives are usually daily means of the individual measurements. To reduce co-location mismatch errors due to the significant difference in horizontal smoothing between S5P and ZSL-DOAS measurements, S5P O3 column values (from afternoon ground pixels at high resolution) are averaged over the footprint of the larger air mass to which the ground-based twilight zenith-sky measurement is sensitive. For more details about the validation methodology, see Lambert et al. (1997, 1999), Balis et al. (2007), Koukouli et al. (2015), Verhoelst et al. (2015), and Garane et al. (2019).

4.2.2 Satellites

S5P TROPOMI L2_O3 total ozone column data have also been compared to Metop-A and Metop-B GOME-2 ozone column data (version GDP 4.8), to Suomi-NPP OMPS-nadir ozone column data, and to S5P ozone column data retrieved with the other S5P operational processor (NRTI vs. OFFL).

4.2.3 Field campaigns and modelling support

Since December 4, 2018, S5P L2_O3 NRTI total ozone data is monitored and assimilated operationally in the Copernicus Atmosphere Monitoring Service system (CAMS), which also assimilates ozone data from a list of other satellite instruments. See Inness et al. (2019) for further details. Specific checks are also carried out by CAMS to verify the effect of a particular event like e.g. a processor upgrade.

4.3 Validation of L2_O3 NRTI

4.3.1 Recommendations for data usage followed

Data users are encouraged to read the Product Readme File (PRF), Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) associated with this data product, all available on https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms.
In order to avoid misinterpretation of the data quality, it is recommended to use only those TROPOMI pixels associated with a qa_value above 0.5. According to validation results, this criterion might be relaxed for data produced with processor versions v01.xx.xx, but nevertheless, caution remains required for qa_value below 0.5. An alternative set of filter criteria for v01.xx.xx L2_O3 NRTI are the following:

- ozone_total_vertical_columnn should be within [0 to 0.45];
- ozone_effective_temperature should be within [180 to 260];
- fitted_root_mean_square should not be larger than 0.01.

From v02.01.03 onwards, the qa_value is redefined in accordance with this alternative set of filter criteria, leading to less data loss when filtering on a qa_value above 0.5, especially at high SZA.

4.3.2 Status of validation

This section presents a summary of the key validation results obtained by the MPC VDAF and by SSP Validation Team (SSPVT) AO projects. This summary is based on coordinated operational validation activities carried out using the Automated Validation Server of the S5P MPC VDAF, the Multi-TASTE versatile multi-platform validation system operated at BIRA-IASB, and the ozone validation system operated at AUTH. This summary takes also into consideration (updates of) the results reported at the SSP First Public Release Validation Workshop (ESA/ESRIN, June 25-26, 2018), at the 3rd SSPVT workshop (ESA/ESRIN, November 11-14, 2019), and at the 4th SSPVT workshop (online meeting hosted by ESA/ESRIN, October 20-21, 2020). Individual contributions to the workshops are archived in https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5p-first-product-release-workshop/sentinel-5p and in https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5p-precursor-workshop-2019/sentinel-5p, while up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal at http://mpc-vdaf.tropomi.eu. The current report also includes quality information provided by the CAMS team at ECMWF, e.g., specific checks carried out during the switch to UPAS processor version 02.01.01 in July 2021.

Current conclusions are valid for the SSP data obtained in the operational phase E2 of the mission, from May 2018 until May 2022, and on the reference data available at the time of this report: typically, until the end of April 2022 for the Dobson and Brewer data, and up to the end of May 2022 for the ZSL-DOAS SAOZ data. For the current report, Brewer and Dobson measurements were obtained through the World Ozone and UV Radiation Data Centre (WOUDC) in Toronto, the NDACC Data Host Facility, and WMO’s Ozone Mapping Centre in Thessaloniki. If a station archives data into both WOUDC and NDACC HDF, the source with the most recent data is adopted. ZSL-DOAS measurements were collected through the SAOZ network Real-Time processing facility operated by CNRS LATMOS (LATMOS_RT). Over the period, with respect to the reference data available at the time of this analysis, of the order of 300 to 2000 co-locations have been identified at about 40 Brewer and Dobson stations and at 12 ZSL-DOAS SAOZ stations, sampling from the Arctic to the Antarctic (Figure 1).
4.3.3 Bias

The systematic difference between S5P L2_O3 NRTI and reference ground-based data at individual stations rarely exceeds 2%, as depicted in Figure 2. The median bias calculated over the entire ground-based networks is of the order of +0.5-1%, S5P reporting higher values than the networks. Between 50˚S and 50˚N, the mean agreement with other satellite data is mostly within 1% as well (Figure 3). This median bias value falls well within the mission requirements (max. bias 5%).

Figure 1: Geographical distribution of Brewer, Dobson and ZSL-DOAS ground-based stations for which co-locations with S5P L2_O3 NRTI ozone data have been used (May 2018 until May 2022).

Figure 2: Meridian dependence of the median (the circular markers) and spread (±1 sigma, the error bars) of the percent relative difference between S5P TROPOMI L2_O3 (PDGS NRTI processor v1.0.0 up to v2.2.1, February 2022) and ground-based (GND) ozone column data, represented at individual stations from the Antarctic to the Arctic and per measurement type (Brewer, Dobson, and ZSL-DOAS). The values in the legend correspond to the median and spread of all median (per station) differences. For clarity, sunrise and sunset ZSL-DOAS results are represented separately (offset by -0.5˚ and +0.5˚ in latitude).
4.3.3.1 Switch to smaller ground pixel size and to UPAS processor v2.1.3 and v2.2.1

On 6 August 2019, the nominal ground pixel resolution of the TROPOMI measurement was reduced to 5.5 x 3.5 km², i.e. shorter by 1.5 km in the along-track direction, by reducing the integration time. Later, on 16 July 2020, the UPAS processor was upgraded to v02.01.03, which includes, among others, improvements in treatment of the surface albedo and in qa_value definition (and consequently, of the filtering applied here). On 5 July 2021, the processor was again upgraded, to v2.2.1, with as most important upgrade the use of the v2 L1b data. Figure 4 shows no evidence of a negative effect on the agreement between satellite and ground-based reference data after either change. Comparisons at Ny-Ålesund and Sodankylä (at the top in Figure 4), stations characterized by a complex and highly variable land/snow/sea/ice albedo, suggest some improvement since UPAS v02.01.03, as expected from replacement of the albedo climatology with the GE_LER albedo. Analysis of a Diagnostic Data Set (not shown here) suggest in increase in overall bias of approximately 0.7%, but in view of the limited size of the DDS, this requires confirmation with a longer record of operational v2.2.1 data. A first analysis of the operational v2.2.1 data (Figure 5) does not reveal a significant change in bias. Feedback from the quality checks performed during the assimilation in CAMS reveals better agreement for the new processor (lower forecast – observation departures) at low solar elevation.

Figure 3: Comparison of the mean percentage differences between three satellite products (S5P TROPOMI L2_O3 NRTI, GOME-2B GDP 4.8, and GOME-2C GDP4.9) and ground-based total ozone data, versus latitude. The Brewer network comparisons are shown in the left-hand panel and the Dobson network comparisons in the right-hand panel. Both ground-based datasets are downloaded from the WOUDC. Time period: is May 2018 – April 2022.

Figure 4: Time series of the difference between the S5P NRTI L2_O3 ozone column data and the SAOZ correlative data at 11 stations from pole to pole, from early 2020 to February 2022. The UPAS processor upgrades to v2.1.3, to v2.2.1, and to 2.3.0 are indicated by the dashed, dotted, and solid lines respectively.
4.3.4 Dispersion

The ±1σ dispersion of the difference (between SSP and reference ground-based network data) around their median value rarely exceeds 3-4% for the comparisons with direct-sun instruments (cf. the error bars depicted in Figure 2). Combining random errors in both satellite and reference measurements with irreducible co-location mismatch effects, it is concluded that the random uncertainty on the SSP measurements falls within the mission requirements of max.±2.5%. Whether the uncertainties reported with the data (the so-called ex-ante uncertainties) are compatible with the observed dispersion can be tested by computing, per station, the $\chi^2$, defined as:

$$\chi^2 = \sum \frac{(SAT - GND - mean(SAT - GND))^2}{\sigma_{SAT}^2 + \sigma_{GND}^2}$$

where, $\sigma_{SAT}$ and $\sigma_{GND}$ refer to the ex-ante satellite and ground-based measurements respectively. The histogram of these $\chi^2$ is shown in Figure 6. The histogram peaks close to one, with only one value near 3. This suggests the ex-ante uncertainties to be realistic.
Figure 6: Histogram of the $\chi^2_r$ values (one per reference instrument) for S5P NRTI total ozone versus different ground-based instruments.

4.3.5 Dependence on influence quantities

The evaluation of potential dependence of the S5P bias and dispersion on the Solar Zenith Angle (SZA, evaluated up to 80°), surface albedo, cloud fraction (CF) and scan sub-index of the TROPOMI measurement does not reveal any variation of the bias larger than 2-3% over the range of these influence quantities (Figure 7).

The scatter of the difference of about 2-4% increases up to 7% at large SZAs and at latitudes beyond 50°, which is expected knowing that random errors in both satellite and reference measurements as well as irreducible co-location mismatch effects increase at high latitudes and low sun elevation.

These results are confirmed by small analysis departures when assimilated in the CAMS system (Inness et al., 2019), albeit with a caveat regarding the effect of surface albedo over snow/ice (e.g., at high latitudes) and a minor systematic effect for TROPOMI ground pixels towards the edges of the swath width (of the order of 1%). See also Section 4.3.7.
4.3.6 Short term variability

Qualitatively, at all of the 50 ground-based monitoring stations, short scale temporal variations in the ozone column as captured by ground-based instruments are reproduced very similarly by S5P, as illustrated in Figure 8. The overall good agreement is corroborated by Pearson correlation coefficients always above 0.95.

Figure 8: Time series of S5P TROPOMI NRTI and Brewer total ozone data at the station of Manchester in the United Kingdom (data courtesy J. Rimmer, University of Manchester).
4.3.7 Geographical patterns

The bias between S5P L2_O3 and other satellite data sets exhibits patterns correlating with weather patterns, atmospheric circulation features, and ground albedo types. When looking at satellite datasets obtained from different satellites (e.g., TROPOMI on S5P in the early afternoon and GOME-2 on Metop-A in the mid-morning), patterns correlating with weather structures and atmospheric circulation might simply reflect – at least partly – real ozone changes between the different satellite overpass times. But patterns correlating with ground albedo types cannot. Furthermore, looking at S5P ozone datasets retrieved from the same Level-1 data processed with different Level-1-to-2 retrieval algorithms, those patterns subsist, as illustrated in Figure 9 where NRTI and OFFL data are compared. These albedo-related differences were largely reduced with the introduction of a dynamic GE_LER albedo in the UPAS v02.01.03 NRTI processor, as can be observed in Figure 10.

Figure 9: Percent relative difference between S5P L2_O3 total ozone data retrieved with the NRTI and OFFL processors (period November 2017 through October 2019). Geographical patterns in this case cannot be associated with real ozone features but reveal rather the effect of using either a surface albedo climatology (NRTI product) or fitting an effective albedo (OFFL product) (courtesy C. Lerot, BIRA-IASB). Note that this comparison no longer holds for data after 16 July 2020 (when the NRTI product changed from the surface albedo climatology to a dynamic GE_LER albedo, in UPAS v02.01.03).

Figure 10: Time evolution of the percent relative zonal mean difference between S5P L2_O3 total ozone data retrieved with the NRTI and OFFL processors (courtesy C. Lerot, BIRA-IASB). It is clear from this figure that the dynamic GE_LER albedo, introduced in the UPAS v02.01.03 NRTI processor (July 2020), solves the large discrepancies between both processors in difficult albedo scenes at high latitudes (70 degrees North and 65 degrees South). Also obvious is the change to the new L1b product in July 2021, affecting more the OFFL columns (with an increase) than those produced by the NRTI processor, as for the former the effective scene albedo depends directly on the radiometric calibration of the spectra, while in the latter the reflectivity is determined by the cloud CAL parameters and the GE_LER surface albedo.

4.3.8 Other features

None to report.
4.4 Validation of L2_O3 OFFL

4.4.1 Recommendations for data usage followed

Data users are encouraged to read the Product Readme File (PRF), Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) associated with this data product, all available on https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms.

In order to avoid misinterpretation of the data quality, it is recommended to use only those TROPOMI pixels associated with a qa_value above 0.5. Nevertheless, it must be noted that at this threshold all data with solar zenith angles larger than 80° are removed (for data with processor version below 02.01.03), leading to a significant rejection of measurements at high latitudes. Validation results suggest that also measurements at larger solar zenith angles are reliable and hence that this cut-off at 80° is not necessary. Consequently, this criterion may be relaxed, nevertheless, caution remains required for qa_value below 0.5. Additional/alternative filter criteria for L2_O3 OFFL are the following:

- ozone_total_vertical_column should range within [0 to 0.45];
- ozone_effective_temperature should range within [180 to 260];
- fitted_root_mean_square should not be larger than 0.01.

From version v02.01.03 onwards, the qa_value is redefined in accordance with these alternative criteria, leading to less data loss at high latitudes when filtering on qa_value above 0.5.

4.4.2 Status of validation

This section presents a summary of the key validation results obtained by the MPC VDAF and by SSP Validation Team (SSPVT) AO projects. This summary is based on coordinated operational validation activities carried out using the Automated Validation Server of the SSP MPC VDAF, the Multi-TASTE versatile multi-platform validation system operated at BIRA-IASB, and the ozone validation system operated at AUTH. This summary takes also into consideration (updates of) the results reported at the SSP First Public Release Validation Workshop (ESA/ESRIN, June 25-26, 2018), at the 3rd SSPVT Workshop (ESA/ESRIN, November 11-14, 2019), and at the 4th SSPVT workshop (online meeting hosted by ESA/ESRIN, October 20-21, 2020). Individual contributions to the workshops are archived in https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5p-first-product-release-workshop/sentinel-5p and in https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5p-precursor-workshop-2019/sentinel-5p, while up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal at http://mpc-vdaf.tropomi.eu.

Current conclusions are valid for the SSP data obtained in the operational phase E2 of the mission, from May 2018 until May 2022, and on the reference data available at the time of this report: typically, until April 2022 for the Dobson and Brewer data, and up to the end of May 2022 for the ZSL-DOAS SAOZ data. For the current report, Brewer and Dobson measurements were obtained through the World Ozone and UV Radiation Data Centre (WOUDC) in Toronto, the NDACC Data Host Facility, and WMO’s Ozone Mapping Centre in Thessaloniki. If a station archives data both into WOUDC and NDACC HDF, the source with the most recent data is adopted. ZSL-DOAS measurements were collected through the SAOZ network Real-Time processing facility operated by CNRS LATMOS (LATMOS_RT). Over the period, with respect to the reference data available at the time of this analysis, of the order of 300 to 2000 co-locations have been identified at about 40 Brewer and Dobson stations and at 12 ZSL-DOAS SAOZ stations, sampling many latitudes from the Arctic to the Antarctic (Figure 11).
4.4.3 Bias

The systematic difference between S5P L2_O3 OFFL and reference ground-based data at individual stations rarely exceeds 2%, as depicted in Figure 12. The median bias calculated over the entire ground-based networks is of the order of +0.3%. Between 50°S and 50°N, the mean agreement with other satellite data derived with the same processor (GODFIT v4) is mostly within 1% as well (Figure 13). This median bias value falls well within the mission requirements (max. bias 5%).
Figure 13: Comparison of the mean percentage differences between three satellite data products (S5P L2_O3 OFFL, OMI GODFIT v4 and OMPS GODFIT v4) and ground-based total ozone data, versus latitude. The Brewer network comparisons are shown in the left-hand panel and the Dobson network comparisons in the right-hand panel. Both datasets are downloaded from the WOUDC. The time period of data used for these plots is May 2018 – April 2022.

4.4.3.1 Switch to smaller ground pixel size and to UPAS processor v2.1.3 and v2.2.1

On 6 August 2019, the nominal ground pixel resolution of the TROPOMI measurement was reduced to 5.5 x 3.5 km², i.e. shorter by 1.5 km in the along-track direction, by reducing the integration time. Later, on 13 July 2020, the UPAS processor was upgraded to v02.01.03, which includes, among others, improvements in qa_value definition (and consequently, in the filtering applied here). On 1 July 2021 the processor was again upgraded, to v2.2.1, with as most important upgrade the use of the v2 L1b data.

Figure 14 shows no evidence of a negative effect on the agreement between satellite and ground-based reference data after either change. Quantifying the improvement of the latest processor upgrade will require a more extensive set of co-locations and is planned for a future quarterly update of this report. Analysis of Diagnostic Data Sets (not shown here) suggests a slight increase of the bias in v2.2.1 by approximately 1.5%, but in view of the limited size of the DDS, this requires confirmation with a longer record of operational v2.2.1 data. A first analysis of the operational v2.2.1 data (Figure 15) confirms this minor increase in bias but without comparisons covering a complete year, neither the DDS nor these first months of operational v2.2.1 can be considered fully representative.

Figure 14: Time series of the difference between the SSP OFFL L2_O3 ozone column data and the SAOZ correlative data at 11 stations from pole to pole, from early 2020 to May 2022. The UPAS processor upgrades to v2.1.3, to v2.2.1, and to v2.3.0 are indicated by the dashed, dotted, and solid lines respectively.
Figure 15: Similar to Figure 12 but for UPAS v2.2.1 data only.

4.4.4 Dispersion

The ±1σ dispersion of the difference (between S5P and reference ground-based network data) around their median value rarely exceeds 3-4% for the comparisons with direct-sun instruments (cf. the error bars depicted in Figure 12). Combining random errors in both satellite and reference measurements with irreducible co-location mismatch effects, it is concluded that the random uncertainty on the S5P measurements falls within the mission requirements of max. 2.5%.

Whether the uncertainties reported with the data (the so-called ex-ante uncertainties) are compatible with the observed dispersion can be tested by computing, per station, the $\chi^2_r$ defined as:

$$\chi^2_r = \frac{1}{N-1} \sum \frac{(\text{SAT}_i - \text{GND}_i - \text{mean(SAT}_i - \text{GND}_i))^2}{\sigma_{\text{SAT}_i}^2 + \sigma_{\text{GND}_i}^2}$$

where, $\sigma_{\text{SAT}_i}$ and $\sigma_{\text{GND}_i}$ refer to the ex-ante satellite and ground-based measurements respectively. The histogram of these $\chi^2_r$ is shown in Figure 16. The histogram peaks close to one, with only a few values above 3. This suggests the ex-ante uncertainties to be realistic.

Figure 16: Histogram of the $\chi^2_r$ values (one per reference instrument) for SSp OFFL total ozone versus different ground-based instruments.
4.4.5 Dependence on influence quantities

The evaluation of potential dependence of the S5P bias and dispersion on the Solar Zenith Angle (SZA, evaluated up to 80°), surface albedo and cloud fraction (CF) of the TROPOMI measurement does not reveal any variation of the bias larger than 2-3% over the range of those influence quantities (Figure 17).

The scatter of the data comparisons of about 2-3% increases up to 5% at large SZAs and at latitudes beyond 50°, which is expected knowing that random errors in both satellite and reference measurements as well as irreducible co-location mismatch errors increase at high latitude and low sun elevation.

Figure 17: Dependence of the difference between S5P OFFL and ground-based Brewer total ozone data on the solar zenith angle (SZA), the surface albedo, the fractional cloud cover, and the scan sub-index of the satellite measurement. Black curve: mean and standard deviation over bins of 10 degrees of SZA, 0.1 of surface albedo, 0.05 of cloud fraction, and 50 pixels over across-track scan.
4.4.6  Short term variability

Qualitatively, at all of the 50 ground-based reference stations, short scale temporal variations in the ozone column as captured by ground-based instruments are reproduced very similarly by S5P, as illustrated in Figure 18. The overall good agreement is corroborated by Pearson correlation coefficients always above 0.95.

Figure 18: Time series of S5P TROPOMI OFFL and Brewer total ozone data at the station of Manchester in the United Kingdom (data courtesy J. Rimmer, University of Manchester).

4.4.7  Geographical patterns

The bias between S5P L2_O3 and other satellite data sets exhibits patterns correlating with weather patterns, atmospheric circulation features, and ground albedo types. When looking at satellite datasets obtained from different satellites (e.g., TROPOMI on S5P in the early afternoon and GOME-2 on Metop-A in the mid-morning), patterns correlating with weather structures and atmospheric circulation might simply reflect – at least partly – real ozone changes between the different satellite overpass times. However, patterns correlating with ground albedo types cannot. Furthermore, looking at S5P ozone datasets retrieved from the same Level-1 data processed with different Level-1-to-2 retrieval algorithms, those patterns subsist, as illustrated in Figure 9 and Figure 10, where NRTI and OFFL data are compared. See Sect. 4.3.7 for more details.

4.4.8  Other features

None to report.
5 Validation Results: L2_O3_TCL

5.1 L2_O3_TCL products and requirements

This Section reports on the validation of the S5P TROPOMI L2_O3_TCL product identified in Table 1. Validation results are discussed with respect to the product quality targets outlined in Table 3.

The S5P O3_TCL data files contain tropospheric ozone columns obtained by the Convective Cloud Differential algorithm (CCD). The CCD data are sampled daily and represent three-day averages of the ozone partial column between surface and 270 hPa (~10.5 km) under cloud-free conditions on a 0.5° latitude by 1° longitude grid between 20°S and 20°N. In contrast to most other S5P products in this document, it concerns a gridded data set, and, it covers about 2/3 of the full vertical range of the tropical troposphere.

Variables related to a second tropospheric ozone algorithm, the Cloud Slicing Algorithm (CSA), are present in the data files but all corresponding entries are set to a fill value for the time being, until further maturation of the algorithm and public release of the CSA product. The CSA data are not discussed in the following.

5.2 Validation approach

Routine validation of the S5P TROPOMI L2_O3_TCL tropospheric ozone data products entails both qualitative, visual inspections of daily maps of product variables, and quantitative comparisons of these to independent reference measurements by ground-based and satellite instruments.

5.2.1 Ground-based networks

Reference measurements by ozonesondes launched at nine stations of the ground-based SHADOZ network (ER_SHADOZ) are compared routinely to S5P data (see Hubert et al., 2021). The SHADOZ data version used here is V06. The ozonesonde profile data are first quality controlled (Hubert et al., 2016, 2021) and then integrated over the vertical range of the S5P CCD product (surface to 270 hPa) to obtain a comparable tropospheric column value. A reference measurement is assumed to be in co-location with a TROPOMI measurement provided that: (a) the SHADOZ station is located in the S5P CCD grid cell, and, (b) the ozonesonde was launched in the satellite time window. Data that do not match these criteria are not used in the calculation of the quality indicators (Figure 21 and Figure 23). If more than one reference tropospheric ozone column falls in a co-location window, then these are averaged prior to comparison. Such a double coincidence occurs very rarely in the considered data sample. Finally, it is important to note that the spatial and temporal sampling properties of satellite and reference data records are quite different, which adds mismatch uncertainties in the comparison results on top of the combined data uncertainties.

5.2.2 Satellites

S5P TROPOMI L2_O3_TCL tropospheric ozone column data are also compared to Aura OMI and Metop-B GOME-2 tropospheric ozone column data using the GODFIT_v4 CCI algorithm developed within ESA’s Climate Change Initiative (CCI). It is based on the GODFIT total column data but the sampling was adapted to allow a more direct comparison to TROPOMI, i.e. 5 days averaging windows instead of monthly data and the tropospheric top pressure set to 270 hPa instead of 200 hPa. The horizontal resolution of the OMI and GOME-2B data products was increased from 1.25° x 2.5° to 1° x 2°.

5.2.3 Field campaigns and modelling support

None for this report.
5.3 Validation of L2_O3_TCL OFFL (CCD)

5.3.1 Recommendations for data usage followed

Data users are encouraged to read the Product Readme File (PRF), Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) associated with this data product, all available on https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms.

In order to avoid misinterpretation of the data quality, we followed the recommendation to use only TROPOMI grid cells associated with a qa_value strictly above 0.7. This screening removes 15.8% of the SSP grid cells, usually between 15-20° latitude in local winter and spring.

5.3.2 Status of validation

This section presents a summary of the key validation results obtained by the MPC VDAF and by SSP Validation Team (SSPV) AO projects. This summary is based on coordinated operational validation activities carried out using the Automated Validation Server of the SSP MPC VDAF and the Multi-TASTE versatile multi-platform validation system operated at BIRA-IASB. This summary takes also into consideration (updates of) the results reported at the S5P L2_O3_TCL and L2_CH4 Data Release Workshop (teleconference, February 20, 2019). Individual contributions to this workshop are archived in https://earth.esa.int/web/sentinel/technical-guides/sentinel-5p/calibration-validation-activities/sentinel-5p-third-products-release-workshop, while up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal at http://mpc-vdaf.tropomi.eu.

Over the period 30 April 2018 – 10 May 2022, the ground-based validation analysis considers 1448 SSP OFFL CCD data products and 974 ozonesonde flights at nine stations across the tropics (Figure 19). SSP data averaged over the entire tropical region are also compared (Figure 23) to GOME-2B data (May 2018 – April 2021) and to OMI data (May 2018 – May 2021).

Figure 19: Median value (left) and half-width of 68% interpercentile (right) of S5P OFFL tropospheric ozone column data (CCD) over the last year of operations (May 2021 – April 2022). Red markers locate the nine ground-based ozonesonde stations used in the validation analysis. Red contours indicate surface elevation (500, 1000, 2000 m). These maps provide context to Figure 20 and Figure 21.

5.3.3 Bias

SSP tropospheric O3 column values are on average larger than the ozonesonde values at all nine stations (Figure 21 and Figure 23). The mean bias over the network is +17% or +3.5 DU (Figure 23, centre and bottom left). This is compliant with the mission requirement for a systematic uncertainty of maximum 25%.
Difference time series between S5P and comparable satellite data (OMI and GOME-2B) averaged over the 20°N – 20°S tropical belt are shown in Figure 22. The agreement with OMI is good, with a mean difference of +0.3 DU or +1.5%. The larger mean difference of +2.5 DU or +15% compared to GOME-2B indicates a slight, general overestimation of TROPOMI which may -at least partly- be attributed to the different overpass times of Metop-B (9:30 descending) and S5P (13:30 ascending) in combination with the diurnal cycle of tropospheric ozone.

The bias between TROPOMI and OMI or GOME_2 (Figure 22) shows a slight positive drift. Relative to the sonde data no drift is observed. If the drift between the instruments is real or caused by the current data availability and the annual cycle has to investigated in the future. The latest update of the Level 1 lead to an increase in the OFFL total column. Up to now, no direct effect of the updated level1 and the respective L2 algorithm adaption is found, there might also be a small increase in the latest part of the time series that might look like a drift.

5.3.4 Dispersion

The half 68% interpercentile of the difference (between S5P and ozonesonde data) ranges within 15-35% or 3.5-8.3 DU (Figure 21 and Figure 23), and the network average is 26% or 4.8 DU (Figure 23, centre and bottom right). Dispersion values at five stations are not compliant with the mission requirement for the random component of the uncertainty (<25%). However, three of these stations are located in an area with large natural percentage variability in the tropospheric O3 field and there is a considerable difference in spatio-temporal sampling between S5P and ozonesonde. In addition, the random component of the uncertainty of the ozonesonde measurement contributes about 5-10% to the observed dispersion in the differences. Hence, the uncertainty of the S5P data is better than the 25% observed dispersion in the comparisons to ozonesonde and therefore overall compliant with the mission requirement.

Satellite-to-satellite comparisons exhibit a dispersion of 2.6-2.9 DU or 14-19% when averaged over the entire tropical belt. This is lower than the mission requirement and the average dispersion in comparisons to the ground-based network (most likely due to the smaller difference in spatio-temporal sampling properties between satellite sensors). Standard deviations shown in Figure 22 (4.2-4.3 DU) are larger since the data are first averaged spatially before the spread is computed. This effectively weighs the result towards regions with higher variability, e.g., the outer tropics.

5.3.5 Dependence on influence quantities

Nothing to report.

5.3.6 Seasonal cycle and shorter term variability

A seasonal cycle appears in the difference between S5P and both other satellite data records from early 2019 onwards (Figure 22). The sampling of the ground-based comparisons is too limited to confirm a seasonal cycle. The phase varies with latitude but, generally, minima are seen around September-January and maxima around March-July. Peak-to-peak amplitude of the cycle lies around 1.5-2.5 DU, depending on latitude and reference instrument (Hubert et al., 2021). In May-August 2020, the difference time series show a brief disruption of the annual cycle. This change in behaviour already starts in April-May and can therefore not be related to the switch in L2 processor version early July 2020.
Signs of an annual cycle in the ground-based comparisons are found at about half of the ozonesonde sites (Figure 24). At Hilo (top left panel), a very clear maximum occurs in boreal winter-spring and a minimum in boreal summer-fall, with a peak-to-peak amplitude of about 16 DU (or 55%). The source of this systematic seasonal effect at Hilo is not well understood. At Paramaribo (top right), up to 10 DU (or 60%) more positive biases are noted during July-November, coinciding with the biomass burning season. More elevated biases of 7-8 DU (or 25-40%) during a few months are seen at other sites around the Atlantic basin as well: Heredia (July-September), Natal (October), Ascension Island (October-November). Whether there is a causal relation between these temporary increases in bias and increased biomass burning around the Atlantic basin during these months is subject of further study.

Co-located S5P and reference measurements correlate fairly well for stations with well-sampled comparison time series. Pearson’s correlation coefficients range between 36% (Paramaribo) and 77% (Natal) at individual stations, while the network average is 63% (Figure 23, top left).

5.3.7 Geographical patterns

Annual median TROPOMI data (May 2021 – April 2022, Figure 19) capture the well-known South Atlantic ozone maximum associated with biomass burning, lightning and ozone precursors, as well as the well-known equatorial Pacific lows. Higher mean levels in the 15°-20° tropical belts are a result of regular intrusions of ozone-rich air from higher latitudes. It shows the ability of S5P to observe the expected large-scale spatial patterns. At smaller scales, however, two sampling-related error patterns are noted. The CCD algorithm requires an ample sampling of input total O3 column data to allow a robust estimate of a reference stratospheric O3 column. This requirement is not always fulfilled and, as a result, random errors of about 1 DU between neighbouring latitude bands are found in many S5P data products. The interplay between cloud coverage and S5P sampling imprints another random error pattern (up to 5 DU) that follows the progression of the S5P orbit. These errors are correlated in time and space and appear at small spatio-temporal scales.

Other known geophysical patterns and oscillations, such as the annual and semi-annual cycles, the biomass-burning season and the Madden-Julian Oscillation, are present in the S5P tropospheric O3 data record as well. An in-depth analysis can be found in Hubert et al. (2021).

5.3.8 Preliminary validation of UPAS processor upgrade to version 02.02.01

The sparse spatio-temporal sampling offered by the SHADOZ network challenges an assessment of the evolution of S5P product data quality. We consider comparisons at four ozonesonde stations with a sufficient coverage of the annual cycle (Hilo, Paramaribo, Ascension Island and Samoa).

Figure 25 shows the mean bias averaged over these stations for each S5P product version. The error bar indicates one standard deviation of the mean estimates. Product versions are separated by Level-1 and Level-2 processor version. The vertical dashed line and shaded area represent the mean over all product versions and its standard error (for the four stations). The latter bias value (3.5±0.9 DU) differs slightly from the results shown in Figure 23 (3.2±0.6 DU) due to the different set of stations included in the average.

There is no statistically significant change in S5P bias due to a change in Level-2 processor (01.01.XX → 02.01.XX), nor due to a change in Level-1 processor (02.01.XX → 02.02.01). The bias difference between 02.02.01 (4.5±1.9 DU) and its predecessor 02.01.XX (3.6±0.9 DU) is not significant. It is not excluded that this is due to the lack of a complete year of data for the intercomparison rather than due to an actual change in S5P bias. Additional S5P and ozonesonde data are needed to further assess the possible impact of the processor change on the data quality.
5.3.9 Other features

CCD data availability is much reduced poleward of ~15° latitude in the winter hemisphere (see e.g. time series at Hilo, Suva or (to a lesser extent) Samoa in Figure 20) since the algorithm requires a sufficient number of highly convective opaque clouds. Most of these are formed in or close to the Intertropical Convergence Zone (ITCZ) located mainly in the summer hemisphere. Suitable cloud conditions therefore occur less frequently in the winter-spring hemisphere.

Filtering on qa_value > 0.7 does not remove all data considered bad. In some S5P products, the screening procedure omits 0.5° latitude bands poleward of 15° latitude in the winter hemisphere which should have been removed. This issue will be tackled in future version of the processor. For the time being, a stricter threshold may solve the issue in some cases.

The change in SSP ground pixel size (on 6 August 2019) does not have a noticeable impact on the quality of S5P tropospheric ozone products (Figure 20 and Figure 21). Estimates of bias and dispersion before and after the change are consistent.

Figure 20: Time series of spatially co-located tropospheric O₃ column data by ozonesonde (red) and by SSP OFFL v01.01.05-v02.03.00 (black). All data were screened following recommendations by the data providers.
Figure 21: Time series of the absolute difference between spatially and temporally co-located S5P and ozonesonde tropospheric $O_3$ column data. The blue line and shaded area shows the median value and the range between the 16% and 84% percentiles. Positive values indicate a high bias of S5P w.r.t. the reference.

Figure 22: Difference time series of daily tropospheric $O_3$ column data averaged over the 20°S – 20°N tropical belt. S5P OFFL CCD data are compared to satellite data by OMI and GOME-2B; positive values indicate a high bias of S5P w.r.t. the reference.
Figure 23: Overview of correlation (top left), median bias (middle & bottom left) and dispersion of the difference (middle & bottom right) of S5P tropospheric O$_3$ column data for each SHADOZ station (black markers). Black vertical bars represent the 68% interpercentile of the difference. The mean, standard error of the mean (1σ) and standard deviation (1σ) of the quality indicator across the network are shown as a horizontal blue line and shaded areas.

Figure 24: Annual cycle of the absolute difference between spatially and temporally co-located S5P and ozonesonde tropospheric O$_3$ column data. Individual comparison pairs are colour-marked by year. Lines indicate the 29-day moving mean of the absolute difference for each year individually (grey) and all years combined (black). The blue line and shaded area shows the median value and the 16-84% percentile range over the entire mission. Positive values indicate a high bias of S5P w.r.t. the reference.
Figure 25: Overview of the evolution of the SSP-ozonesonde bias averaged over four ozone stations with sufficient coverage of the annual cycle (Hilo, Paramaribo, Ascension Island and Samoa). The error bar shows one standard deviation of the mean estimates. Product versions are separated by Level-1 and Level-2 processor version. The vertical dashed line and shaded area represent the mean over all product versions and its standard error (for the four stations). Co-location sample size is displayed on the left.
6 Validation Results: L2_O3_PR

6.1 L2_O3_PR products and requirements

This Section reports on the validation of the S5P TROPOMI L2_O3_PR data product identified in Table 1. Validation results are discussed with respect to the product quality targets outlined in Table 3. These results are obtained by the MPC VDAF and by the S5P Validation Team (S5PVT) AO project CHEOPS-5p. The operational validation activities are carried out using the Automated Validation Server of the MPC VDAF and the Multi-TASTE versatile multi-platform validation system operated at BIRA-IASB. This section provides validation results for the near-real-time (NRT) and offline (OFFL) PDGS operational processing streams (here up to May 2022), which are available on ESA’s data hub since November 17 and November 16, 2021, respectively. Due to a change in the processing timeout, NRT data are publicly available starting from November 24, 2021, only. The operational S5P L2_O3_PR orbit data files contain, for each individual observation, the ozone number density on 33 pressure levels as retrieved by KNMI’s operational algorithm v02.03.01 (from L1B v02.00.00), the integrated tropospheric and total ozone columns, and six integrated sub-columns (0-6, 6-12, 12-18, 18-24, 24-32, and 32-82 km). For the validation activities presented here, the station overpass files obtained from the PDGS processor in HARP format (v1.14) are considered.

6.2 Validation approach

Validation of the S5P TROPOMI L2_O3_PR ozone profile data entails quantitative comparisons to independent reference measurements collected from ground-based monitoring networks, cross-validation with other satellite instruments, assessment of the retrieved information content based on the analysis of the associated averaging kernels, and visual inspections of daily maps of S5P ozone data and associated parameters.

6.2.1 Ground-based networks

S5P TROPOMI L2_O3_PR ozone profile data are compared to ground-based measurements acquired by instruments contributing to WMO’s Global Atmosphere Watch (GAW), the Network for the Detection of Atmospheric Composition Change (NDACC), Southern Additional Ozonesonde programme (SHADOZ), and Tropospheric Ozone Lidar Network (TOLNET): (1) balloon-borne ozonesondes, (2) stratospheric differential absorption ozone lidars (DIAL), and (3) tropospheric DIAL. The ground-based data are collected through ESA’s Validation Data Centre (EVDC) and the respective data host facilities of the ground-based networks.

6.2.1.1 Balloon-borne ozonesonde

Launched on board of small meteorological balloons, electrochemical ozonesondes measure the vertical distribution of atmospheric ozone partial pressure from the ground up to burst point, typically around 30 km. Their estimated bias is smaller than 5 %, and the precision remains within the order of 3 % (Smit et al., 2007). Caveats for using ozonesonde datasets include errors depending on instrument set up (buffer solution, pump efficiency correction...) and changes in ozonesonde type and measurement parameters with time. In the framework of the MPC, the VDAF-AVS performs automated data comparisons with ozonesonde datasets collected through the EVDC. These data originate from the NDACC Data Host facility, the SHADOZ archive, and World Ozone and UV Data Centre (WOUDC).
6.2.1.2 Differential absorption ozone lidars

Ground-based differential absorption lidars (DIAL) measure either the tropospheric ozone profile or the stratospheric ozone profile. They perform network operation in the framework of the global NDACC and the US TOLNET (Tropospheric Ozone Lidar Network) networks. Stratospheric ozone lidars measure the ozone number density from above the tropopause up to about 45–50 km altitude with a vertical resolution that declines with altitude from 0.3 to 3–5 km. The estimated uncertainty due to systematic and random effects is about 2 % between 20 and 35 km and increases to 10 % outside this altitude range where the signal-to-noise ratio is smaller (Keckhut et al., 2004). With respect to ozonesondes, tropospheric ozone lidar measurements between 3 km and 10 km show a mean difference below 2 % and a root-mean-square deviation below 3 %, which are well within the combined uncertainties of the two measurement techniques (Leblanc et al., 2018). The MPC VDAF and its AVS make use of DIAL ozone profile data available through EVDC, originating from the NDACC Data Host facility, and the TOLNET data archive.

6.2.2 Satellite intercomparisons

Comparison to other satellite data extend ground-based validation to the global domain and increase the number of data comparisons. For stratospheric ozone, comparisons to limb and solar occultation sounders (MLS, OMPS-limb, ACE-FTS) are appropriate. For tropospheric ozone, comparisons can be made to OMI and OMPS-nadir, where the OMPS-nadir measurements have the best spatial and temporal co-registration with TROPOMI.

6.2.3 Analysis of information content

The information content of the S5P ozone profile data is assessed through algebraic analysis of the associated averaging kernel (AK) matrix generated by the same S5P processing algorithm. The row sums of the AK matrix indicate the vertical sensitivity of the S5P ozone profile retrieval (Rodgers, 2000). The trace of the AK kernel matrix gives the Degree of Freedom of the Signal (DFS), to be understood here as the amount of vertical sub-columns with independent ozone information from each other. The Full Width at Half Maximum (FWHM) of the AK corresponding to a given altitude gives an indication of the effective vertical resolution of the retrieved profile at this altitude. This effective resolution of the retrieved information is not the numerical resolution of the vertical grid used for the retrieval process, which is usually much higher than the true, physical resolution of the retrieved information. The true altitude registration of the retrieved profile information at a given altitude of calculation can be estimated as the barycentre or peak position of the associated AK at this calculation altitude.

6.2.4 Analysis of daily global maps

The MPC VDAF-AVS creates daily global maps of the six partial columns provided in the ozone profile product, together with the integrated total column. The latter is compared with the daily global map of the TROPOMI total column retrieval to assess their mutual consistency. Daily global maps easily allow identifying data gaps, retrieval artefacts, along-orbit striping, and other large-scale features that are not typically detected through comparison with respect to point-like ground-based data.

6.2.5 Parameter correlation checks

Using the in-house PyCAMA software, correlation checks are performed by KNMI on a broad selection of satellite data parameters within the orbit files. These checks provide a view on single-orbit features, correlations between retrievals of subsequent pixels, the appropriateness of the data flagging, etc. Relevant results can be found on the TROPOMI Portal for Level-2 Quality Control [ER_L2QC].
6.2.6  Field campaigns and modelling support

No specific campaigns and model-based studies have been foreseen to support the validation of SSP TROPOMI ozone profile data. However, ozone profile measurements acquired during campaigns not specific to TROPOMI validation will be considered as well (e.g. the ozonesonde MATCH campaigns).

6.3  Validation of L2_O3_PR NRTI

6.3.1  Recommendations for data usage followed

Data users are encouraged to read the Product Readme File (PRF), Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) of this data product, available online through the following link: https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-5p/products-algorithms. In order to avoid misinterpretation of the data quality, we follow the PUM recommendation to use only TROPOMI ozone profile retrievals associated with a qa_value above 0.5.

6.3.2  Status of validation

Comparison results between ground-based measurements and coincident TROPOMI L2_O3_PR pixels (closest pixel on the same day with qa_value > 0.5) are obtained through the versatile Multi-TASTE validation system at BIRA-IASB, as part of both MPC and S5PVT CHEOPS-5p validation activities. Prior to their comparison to S5P data, ground-based measurements – acquired at higher vertical resolution than S5P profile data – are convolved with the averaging kernels associated with the S5P retrievals to account for vertical smoothing differences (see e.g. Rodgers and Connor, 2003, Calisesi et al. 2005, Keppens et al., 2019). In Figure 27, the difference of L2_O3_PR ozone number density values with respect to reference measurements is reported as a function of a selection of influence quantities (colour scales). Also included are the level-specific chi-square tests (von Clarmann, 2006; Keppens et al., 2019) and a selection of information content diagnostics: vertical sensitivity, altitude registration offset, and averaging kernel full width at half maximum (FWHM). The geographical distribution of the FRM stations depicted in Figure 26 indicates the domain of applicability of the validation results.

For the routine validation of the S5P/TROPOMI ozone profiles, the automated validation server (AVS, http://mpc-vdaf-server.tropomi.eu/o3-profile) deployed within the MPC VDAF facility collects S5P ozone profile data and correlative measurements to identify suitable co-locations, compare the co-located data, and produce SSP data quality indicators. The VDAF-AVS produces curtain plots (ozone number density as a function of altitude and time) of the satellite data at a selection of ground-based ozonesonde stations, together with curtain plots showing the difference between S5P and ground-based data. The VDAF-AVS also provides statistical estimates of the bias and dispersion of S5P data with respect to the ground-based measurements.
Figure 26: Geographical distribution of the ozonesonde and lidar stations with which co-locations with SSP L2_O3_PR NRTI data have been identified and used in the data comparisons reported hereafter.

Figure 27: Comparison between SSP L2_O3_PR NRTI ozone number density profile data and all co-located ground-based reference measurements. Every plate shows six graphs, respectively, from left to right: the difference and the percent relative difference between SSP and ozonesonde (left panels) or lidar (right panels) data, the chi-square profile, the vertical sensitivity, the altitude registration offset, and the averaging kernel FWHM associated with the SSP retrieval. The colour scale indicates the DFS (upper panels), and the TROPOMI solar zenith angle (lower panels). Black dashed lines show mean values (thick lines) and standard deviations (thin lines, around the mean), while grey dashed lines indicate the mean difference between the a-priori profile and the reference measurement. Dotted black lines indicate the total ex-ante (inductive) uncertainty of TROPOMI and the reference measurements combined (around the mean difference).
6.3.3 Vertical sensitivity, resolution and registration

The information content of the S5P ozone profile data is assessed through the analysis of the associated averaging kernels generated by the same L2_O3_PR data processor, as described in Section 6.2.3.

The vertical sensitivity of the S5P ozone profile data is nearly equal to unity at altitudes from about 20 km up to 50 km. It decreases rapidly at altitudes below 10 km and above 50 km. Around the tropopause between 10 and 20 km an oversensitivity (larger than one) can be observed, which is a rather typical compensating effect for the undersensitivity below in nadir profile retrievals.

The retrieved information on ozone is distributed into about six vertical sub-columns of independent information. The effective vertical resolution of the profile retrieval usually ranges within 10-15 km, with a minimum close to 7 km in the middle stratosphere (around 30-40 km).

The altitude registration of the retrieved profile information usually is close to the nominal retrieval altitude in the 20-50 km altitude range, and shows positive and negative offsets of up to 10 km below and above the 20-50 km altitude range, respectively.

6.3.4 Bias

Compared to ozonesonde data, S5P L2_O3_PR NRTI data has a bias below about 5-10 % in the troposphere up to the UTLS (upper troposphere to lower stratosphere). Tropospheric lidar data show a positive tropospheric bias up to 30 %, but this is for a single station only (Table Mountain, USA).

In the stratosphere, data comparisons conclude to a bias of maximum 5-10 % but with slight vertical oscillations (positive and negative). These oscillations of the bias may be due to a typically larger a-priori error in the mid and high stratosphere (above 20 %) in comparison with other retrievals.

Positive biases that occur above 30 km for ozonesondes and above about 45 km for lidars should not be considered, as the reference instruments become less trustworthy above those altitudes.

6.3.5 Dispersion

S5P data comparisons with ozonesonde and lidar data show a dispersion of order of 20 % in the troposphere, UTLS, and upper stratosphere, and below 10 % in the stratosphere in between.

6.3.6 Chi-square tests

Chi-square tests ($\chi^2 = (\Delta x)^T S^{-1} \Delta x$) allow verifying whether the observed differences $\Delta x$ between the satellite and reference profiles are consistent with the ex-ante (predicted) uncertainties on the difference $S_{\Delta}$ (von Clarmann, 2006; Keppens et al., 2019). The latter contains the satellite and reference covariances, and uncertainties that are due to sampling, smoothing, and retrieval differences. By application of vertical averaging kernel smoothing, however, retrieval differences and vertical sampling and smoothing differences are removed from the difference covariance $S_{\Delta}$ (Keppens et al., 2019). This means that for the results presented here the difference covariance mainly contains the satellite and reference covariances. Horizontal and temporal sampling and smoothing differences are minimized to a nearly negligible level by application of strict collocation criteria (reference station within satellite pixel and a few hours measurement time difference only).
The chi-square plots in Figure 27 demonstrate that on average the observed differences confirm ($\chi^2 \leq 1$) the ex-ante satellite and ground uncertainty estimates in the stratosphere (below the levels where a positive bias occurs due to reference instrument degradation). However, around the tropopause and below (around 15-20 km and lower), the mean chi-square value increases up to about two. Here, the predicted (random) satellite uncertainty is smaller than what is actually observed (assuming correct reference uncertainties). This can also be seen in the difference plots, as the thin dashed lines representing dispersions of the difference are further away from the mean difference than the dotted lines representing combined ex-ante uncertainties.

6.3.7 Dependence on influence quantities

The amount of SSP data that is currently available to the validation is too limited to enable the proper detection of any dependence on the following influence quantities and parameters: tropospheric column, fractional cloud cover, DFS, latitude, qa_value, surface albedo, SZA, scan angle. For now, increased sensitivities and higher effective vertical resolutions (FWHM) can be observed for high solar zenith angles, as can be expected. The increased stratospheric bias oscillations that correspondingly occur (now most clearly from the lidar comparisons) require further study. On the other hand, a reduced DFS, sensitivity, and retrieval quality are sometimes observed for scenes that have both high SZA and high surface albedo, especially over the sea around the Antarctic. This feature is currently being examined. A detailed assessment will be possible in the future when more operational S5P ozone profile data become available.

6.3.8 Short term variability

Same as Sub-section 6.3.7: To be addressed when more data become available.

6.3.9 Geographical patterns

Global maps for February 19, 2022, are shown in Figure 28. Slight along-orbit striping can be observed, especially in the middle stratosphere (24-32 km sub-column).

Global maps of the integrated L2_O3_PR ozone profile data and of the L2_O3 total column data (bottom row) look mutually consistent (same colour scale).

6.3.10 Other features

Orbit curtain plots reveal that for some TROPOMI pixels, particularly in the beginning of the orbit, the retrieved ozone profiles deviate strongly and non-physically from the a-priori. This issue needs further examination and might require an update of the data flagging (in terms of qa_value definition) in the future.

In the absence of clouds, data files sometimes contain negative surface albedo values. The TROPOMI ground pixels affected by this anomaly are usually located at the east and west edges of the across-orbit measurement swath. For now, these negative values are set to zero in the radiative transfer code and validation tools (as an influence quantity, also see above). The same could be done in the ozone profile data distribution to users.
Figure 28: Upper six panels: daily global map for the six partial columns in the S5P L2_O3_PR NRTI v02.03.01 ozone profile product of April 29, 2022, as a function of altitude. The two lower maps show the map of total ozone column values obtained by integration of the L2_O3_PR profile data (left), and the map of L2_O3 total ozone column values for the same day (right) to check for their mutual consistency.
6.4 Validation of L2_O3_PR OFFL

6.4.1 Recommendations for data usage followed

Data users are encouraged to read the Product Readme File (PRF), Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) of this data product, available online through the following link: https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-5p/products-algorithms. In order to avoid misinterpretation of the data quality, we follow the PUM recommendation to use only TROPOMI ozone profile retrievals associated with qa_value above 0.5.

6.4.2 Status of validation

Comparison results between ground-based reference measurements and coincident TROPOMI L2_O3_PR pixels (closest pixel on the same day with qa_value > 0.5) are obtained through the versatile Multi-TASTE validation system at BIRA-IASB, as part of both MPC and S5PVT CHEOPS-5p validation activities. Prior to their comparison to S5P data, ground-based measurements – acquired at higher vertical resolution than S5P profile data – are convolved with the averaging kernels associated with the S5P retrievals to account for vertical smoothing differences (see e.g. Rodgers and Connor, 2003, Calisesi et al. 2005, Keppens et al., 2019). In Figure 29, the difference of L2_O3_PR ozone number density values with respect to reference measurements is reported as a function of a selection of influence quantities (colour scales). Also included are the level-specific chi-square tests (von Clarmann, 2006; Keppens et al., 2019) and a selection of information content diagnostics: vertical sensitivity, altitude registration offset, and averaging kernel full width at half maximum (FWHM). The geographical distribution of the FRM stations depicted in Figure 26 indicates the domain of applicability of the validation results.

For the routine validation of the S5P/TROPOMI ozone profiles, the automated validation server (AVS, http://mpc-vdaf-server.tropomi.eu/o3-profile) deployed within the MPC VDAF facility collects S5P ozone profile data and correlative measurements to identify suitable co-locations, compare the co-located, data and produce S5P data quality indicators. The VDAF-AVS produces curtain plots (ozone number density as a function of altitude and time) of the satellite data at a selection of ground-based ozonesonde stations, together with curtain plots showing the difference between S5P and ground-based data. The VDAF-AVS also provides statistical estimates of the bias and dispersion of S5P data with respect to the ground-based measurements.
Figure 29: Comparison between SSP L2_O3_PR OFFL ozone number density profile data and all co-located ground-based reference measurements. Every plate shows six graphs, respectively, from left to right: the difference and the percent relative difference between SSP and ozonesonde (left panels) or lidar (right panels) data, the chi-square profile, the vertical sensitivity, the altitude registration offset, and the averaging kernel FWHM associated with the SSP retrieval. The colour scale indicates the DFS (upper panels), and the TROPOMI solar zenith angle (lower panels). Black dashed lines show mean values (thick lines) and standard deviations (thin lines, around the mean), while grey dashed lines indicate the mean difference between the a-priori profile and the reference measurement. Dotted black lines indicate the total ex-ante (inductive) uncertainty of TROPOMI and the reference measurements combined (around the mean difference).

6.4.3 Vertical sensitivity, resolution and registration

The information content of the SSP ozone profile data is assessed through the analysis of the associated averaging kernels generated by the same L2_O3_PR data processor, as described in Section 6.2.3.

The vertical sensitivity of the SSP ozone profile data is nearly equal to unity at altitudes from about 20 km up to 50 km. It decreases rapidly at altitudes below 10 km and above 50 km. Around the tropopause between 10 and 20 km an oversensitivity (larger than one) can be observed, which is a rather typical compensating effect for the under-sensitivity below in nadir profile retrievals.

The retrieved information on ozone is distributed into about six vertical sub-columns of independent information. The effective vertical resolution of the profile retrieval usually ranges within 10-15 km, with a minimum close to 7 km in the middle stratosphere (around 30-40 km).

The altitude registration of the retrieved profile information usually is close to the nominal retrieval altitude in the 20-50 km altitude range, and shows positive and negative offsets of up to 10 km below and above the 20-50 km altitude range, respectively.

6.4.4 Bias

Compared to ozonesonde data, SSP L2_O3_PR OFFL data has a bias below about 5-10 % in the troposphere up to the UTLS (upper troposphere to lower stratosphere). Tropospheric lidar data show a positive tropospheric bias up to 30 %, but this is for a single station only (Table Mountain, CA).
In the stratosphere, data comparisons conclude to a bias of maximum 5-10% but with slight vertical oscillations (positive and negative). These oscillations of the bias may be due to a typically larger a-priori error in the mid and high stratosphere (above 20%) in comparison with other retrievals.

Positive biases that occur above 30 km for ozonesondes and above about 45 km for lidars should not be considered, as the reference instruments become less trustworthy above those altitudes.

### 6.4.5 Dispersion

SSP data comparisons with ozonesonde and lidar data show a dispersion of order of 20% in the troposphere, UTLS, and upper stratosphere, and below 10% in the stratosphere in between.

### 6.4.6 Chi-square tests

Chi-square tests \( \chi^2 = (\Delta x)^T S_{\Delta}^{-1} \Delta x \) allow verifying whether the observed differences \( \Delta x \) between the satellite and reference profiles are consistent with the ex-ante (predicted) uncertainties on the difference \( S_{\Delta} \) (von Clarmann, 2006; Keppens et al., 2019). The latter contains the satellite and reference covariances, and uncertainties that are due to sampling, smoothing, and retrieval differences. By application of vertical averaging kernel smoothing, however, retrieval differences and vertical sampling and smoothing differences are removed from the difference covariance \( S_{\Delta} \) (Keppens et al., 2019). This means that for the results presented here the difference covariance mainly contains the satellite and reference covariances. Horizontal and temporal sampling and smoothing differences are minimized to a nearly negligible level by application of strict collocation criteria (reference station within satellite pixel and a few hours measurement time difference only).

The chi-square plots in Figure 29 demonstrate that on average the observed differences confirm \( \chi^2 \leq 1 \) the ex-ante satellite and ground uncertainty estimates in the stratosphere (below the levels where a positive bias occurs due to reference instrument degradation). However, around the tropopause and below (around 15-20 km and lower), the mean chi-square value increases up to about two. Here, the predicted (random) satellite uncertainty is smaller than what is actually observed (assuming correct reference uncertainties). This can also be seen in the difference plots, as the thin dashed lines representing dispersions of the difference are further away from the mean difference than the dotted lines representing combined ex-ante uncertainties.

### 6.4.7 Dependence on influence quantities

The amount of SSP data that is currently available to the validation is too limited to enable the proper detection of any dependence on the following influence quantities and parameters: tropospheric column, fractional cloud cover, DFS, latitude, qa_value, surface albedo, SZA, scan angle. For now, increased sensitivities and higher effective vertical resolutions (FWHM) can be observed for high solar zenith angles, as can be expected. The increased stratospheric bias oscillations that correspondingly occur (now most clearly from the lidar comparisons) require further study. On the other hand, a reduced DFS, sensitivity, and retrieval quality are sometimes observed for scenes that have both high SZA and high surface albedo, especially over the sea around the Antarctic. This feature is currently being examined. A detailed assessment will be possible in the future when more operational SSP ozone profile data become available.

### 6.4.8 Short term variability

A detailed assessment will be possible in the future when more operational SSP ozone profile data become available.
6.4.9 Geographical patterns

Global maps for February 18, 2022 are shown in Figure 30. Slight along-orbit striping can be observed, especially in the middle stratosphere (24-32 km sub-column).

Global maps of the integrated L2_O3_PR ozone profile data and of the L2_O3 total column data (bottom row) look mutually consistent (same colour scale).

6.4.10 Other features

Orbit curtain plots reveal that for some TROPOMI pixels, particularly in the beginning of the orbit, the retrieved ozone profiles deviate strongly and non-physically from the a-priori. This issue needs further examination and might require an update of the data flagging (in terms of qa_value definition) in the future.

In the absence of clouds, data files sometimes contain negative surface albedo values. The TROPOMI ground pixels affected by this anomaly are usually located at the east and west edges of the across-orbit measurement swath. For now, these negative values are set to zero in the radiative transfer code and validation tools (as an influence quantity, also see above). The same could be done in the ozone profile data distribution to users.
Figure 30: Upper six panels: daily global map for the six partial columns in the S5P L2_O3_PR OFFL v02.03.01 ozone profile product of April 29, 2022, as a function of altitude. The two lower maps show the total ozone column values obtained by integration of the L2_O3_PR profile data (left), and the map of L2_O3 total ozone column values for the same day (right) to check for their mutual consistency.
7 Validation Results: L2_NO2

7.1 L2_NO2 products and requirements

This section reports on the validation of the following geophysical variables of the S5P TROPOMI L2_NO2 data products identified in Table 1: the NO$_2$ tropospheric column, the stratospheric NO$_2$ column, and the NO$_2$ total column. Validation results are discussed with respect to the product quality targets outlined in Table 3. The operational (E2) phase for the S5P TROPOMI mission starts with orbit #02818 on 2018/04/30.

The OFFL data has been reprocessed to version 01.02.02 for the time span 2018/05/01 to 2018/10/17. After that, the processor version has been changed eight times from 01.02.00 up to the current version 02.03.01 started on 2021/11/14. The NRTI data covers the full range of versions from 01.00.01 to 02.03.01 as there is no reprocessing necessary. Due to important upgrades in v01.04.00 and 02.02.00, a scientific reprocessing has already been performed (SSP-PAL, Product Algorithm Laboratory). OFFL and NRTI products may differ because they are retrieved using ECWMF forecast meteorological data with a time difference as input for the CTM. Subsection 7.4 demonstrates evidence that NRTI and OFFL data do not differ significantly and that their respective validations yield similar conclusions.

7.2 Validation approach

7.2.1 Ground-based monitoring networks

Tropospheric NO$_2$ – MAX-DOAS UV-Visible Spectrometers

S5P TROPOMI L2_NO2 tropospheric nitrogen dioxide column data are routinely compared to reference measurements acquired by MAX-DOAS (Multi-Axis Differential Optical Absorption Spectroscopy) UV-Visible spectrometers. Several of those instruments perform network operation in the context of the Network for the Detection of Atmospheric Composition Change (NDACC). MAX-DOAS tropospheric NO$_2$ column data have a maximum bias of 20% and a precision better than 30% at this set of stations. The validation with MAX-DOAS data from Nitrogen Dioxide and FORmaldehyde VALidation (NIDFORVAL) has been included in a harmonized validation effort (Verhoelst et al., 2021). Since then, contact with NIDFORVAL PIs to extend their datasets and provision on ESA Atmospheric Validation Data Centre (EVDC) through conversion to fully GEOMS (Generic Earth Observation Metadata Standard) and HARP compatible data lead to the inclusion of new stations to the VDAF-AVS and several time-period updates. Careful revision (and eventual reprocessing) are ongoing for the Thessaloniki stations and the UNAM stations (Cuautitlan, Unam and Vallejo).

Stratospheric NO$_2$ – ZSL-DOAS UV-Visible Spectrometers

S5P TROPOMI L2_NO2 stratospheric nitrogen dioxide column data are compared routinely to reference measurements acquired by Zenith-Scattered Light Differential Optical Absorption Spectroscopy (ZSL-DOAS) UV-Visible spectrometers (Pommereau and Goutail, 1988; Hendrick et al., 2011). The instruments perform network operations in the context of the Network for the Detection of Atmospheric Composition Change (NDACC). The ZSL-DOAS validation data (VDAF-AVS) have been obtained through the SAOZ near-real-time processing facility operated by the CNRS LATMOS (see Figure 31, red dots). They are complemented with measurements from 13 other NDACC affiliated ZSL-DOAS instruments (blue and green dots). The stations are located between 79°N and 75°S.
NDACC field intercomparison campaigns (Roscoe et al., 1999; Vandaele et al., 2005) conclude to an uncertainty of about 4-7% on the slant column density. Converting the slant column into a vertical column using a zenith-sky AMF, the uncertainty on the vertical column is estimated to be on the order of 10-14% for the latest data processing version (Yela et al., 2017; Bognar et al., 2019). A limiting factor comes from the temperature dependence of the NO$_2$ absorption cross-sections used in the DOAS retrieval of the slant column density. Most of the NDACC instruments use cross-sections at a single temperature of 220 K, which introduces a seasonal error of up to a few percent at middle and high latitudes.

![Figure 31: Geographical distribution of the NDACC ZSL-DOAS instruments routinely measuring stratospheric NO$_2$ and yielding co-locations with the current S5P L2_NO2 datasets. Stations marked with a red dot contribute fast delivery data coming from the LATMOS_RT facility. Blue and green dots depict the NDACC stations that contributes ZSL-DOAS data directly through the NDACC DHF and the AO project NIDFORVAL, respectively.](image)

To account for effects of the photochemical diurnal cycle of stratospheric NO$_2$, the ZSL-DOAS measurements which are obtained two times a day at twilight are adjusted to the S5P overpass time using a model-based factor. This is calculated with the PSCBOX 1D stacked-box photochemical model (Errera and Fonteyn, 2001; Hendrick et al., 2004), initiated with daily fields from the SLIMCAT chemistry-transport model (CTM). The amplitude of the adjustment depends strongly on the effective SZA assigned to the ZSL-DOAS measurements which is here taken to be 89°. The uncertainty related to this adjustment is in the order of 10%. To reduce mismatch errors due to the significant horizontal smoothing differences between S5P and ZSL-DOAS measurements, S5P NO$_2$ values (from ground pixels at high resolution) are averaged over the air mass footprint where ground-based zenith-sky measurements are sensitive. Additional confirmation is obtained by comparison with 3 mountain-top PGN (Pandonia Global Network) instruments where the measured signal corresponds more to the S5P L2_NO2 stratospheric column rather than the total column. These are Altzomoni (3985 m), Izaña (2360 m), and Mauna Loa (4169 m).

**Stratospheric NO$_2$ – FTIR spectrometers**

The ground-based FTIR instruments measure stratospheric NO$_2$ (e.g. Hendrick et al., 2012, Bognard et al., 2019, Garcia et al., 2021) with a precision of about 8-12% and a systematic error of about 10%. Within the AO project NIDFORVAL, the retrieval settings have been harmonized and applied for 23 FTIR stations for S5P validation. We build the collocated pairs to be compared in several steps:
When the collocation is not above the FTIR station, we use the line-of-sight of the instrument (FTIR are direct sun measurements) instead and calculate the position along the line-of-sight corresponding to the altitude where the NO$_2$ FTIR averaging kernels show the maximum sensitivity (~30-35km). Then, S5P pixels are selected within 50 km of this position (about 150-200 pixels). Only pixels with qa_value > 0.5 are used. A collocation pair is only kept if at least 10 pixels can be averaged.

- The time coincidence criterion is set to ±1 hour of the satellite overpass time.
- The comparison methodology is the same as for HCHO validation using FTIR data (Vigouroux et al., 2020): (i) The FTIR a priori profile is substituted with the TROPOMI L2_NO2 one to get a corrected FTIR profile. (ii) The corrected profile is smoothed with the TROPOMI averaging kernel (Rodgers and Connor, 2003). In this process, since the TROPOMI averaging kernels are zero below the tropopause for the stratospheric NO$_2$, the tropospheric part of the FTIR profile is removed, and only stratospheric columns from both products are indeed compared. (iii) Both the individual manipulated FTIR columns and the individual S5P manipulated pixel columns are then averaged.

- Finally, the relative median bias at a single station is estimated by the median relative difference: \(\text{Med}[(\text{SAT}-\text{REF})/\text{REF}]\). Absolute-scale dispersion is estimated by the scaled median absolute deviation from the median (MAD): \(1.4826\times\text{MEDIAN}[\text{ABS}(\text{DIFF}-\text{MEDIAN}(\text{DIFF}))]\). The scaling factor of 1.4826 ensures that for a normal distribution, the MAD is equal to the standard deviation.

**Total NO$_2$ – Pandora Direct-Sun UV-Visible Spectrometers**

TROPOMI L2_NO2 nitrogen dioxide summed column data (troposphere + stratosphere) are routinely compared to reference measurements acquired by Pandora instruments. They perform network operation in the context of the Pandonia Global Network (PGN). Pandora total NO$_2$ data have maximum bias of 10-15% and a precision of roughly 0.28 Pmolec/cm$^2$ (about 10%). The comparison criteria on the VDAF Automated Validation Server are: TROPOMI L2_NO2 data with qa_value > 0.5; the TROPOMI ground pixel contains the Pandora station; Pandora negative values are excluded and only measurements with a flag not equal to 0 and 10 are used; and the Pandora measurement closest in time is selected, with a maximum time difference of 30 min. If the Pandora instrument operates at an elevated station above low-lying tropospheric pollution, the Pandora measurement in absence of free troposphere NO$_2$ can also be representative of the stratospheric NO$_2$ column.

**7.2.2 Satellites**

TROPOMI L2_NO2 nitrogen dioxide column data are also compared to data from the Ozone Monitoring Instrument (OMI) retrieved with both the QA4ECV and the IUP-UB algorithm. OMI is on board the EOS-Aura satellite that was launched in July 2004.

**7.2.3 Field campaigns and modelling support**

None for this report.
7.3 Validation of L2_NO2

7.3.1 Recommendations for data usage

In order to avoid misinterpretation of the data quality, it is recommended to only use TROPOMI NO$_2$ measurements with $qa\_value \geq 0.75$. This removes cloudy scenes (cloud radiance fraction $> 0.5$), scenes covered by snow/ice, several other errors, and problematic retrievals. For stratospheric NO$_2$ retrievals and data comparisons, clouds are less of a problem, so that $qa\_value \geq 0.5$ can be used. Data users are encouraged to read the Product Readme File (PRF), Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) associated with this data product, all available on https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms.

7.3.2 Status of validation

This section presents a summary of the key validation results obtained by the MPC VDAF and by SSPVT AO projects. Routine operations validation activities rely on the Automated Validation Server of the MPC VDAF, the Multi-TASTE versatile multi-platform validation system operated at BIRA-IASB, the validation tools of IUP-B, and the HARP toolset (version 1.6).

An in-depth discussion of the routine validation results up to March 2020 by Verhoelst et al. (2021) was published. Improvements of V2.2 are discussed by van Geffen et al. (2022). The reprocessed SSP-PAL data of V2.3.1 is verified by Eskes et al. (2021).

Up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal. The PGN steadily extends its Pandora data provision on EVDC, making it accessible for the VDAF Automated Validation Server. Furthermore, there is regular contact with NIDFORVAL PIs to extend their MAX-DOAS datasets and their provision on EVDC (through conversion to fully GEOMS and HARP compatible data) and to do time-period updates. Moreover, contacts with new MAX-DOAS station PIs have been established that lead to new stations inclusions. Within NIDFORVAL, the FTIR PIs applied the harmonized settings and submit their data individually to the NIDFORVAL’s PI. In summary, the FRM data streams cover the time period from May 2018 to May 2022:

- MAX-DOAS tropospheric columns from the NIDFORVAL AO project from 27 stations. Vertical smoothing harmonization could be applied for 14 stations. MAX-DOAS tropospheric column data at the VDAF Automated Validation Server via EVDC are available from 8 stations (De Bilt, Cabauw (KNMI), Uccle, Xianghe (BIRA-IASB), Bremen, Athens (IUP-B), Mainz (MPIC); and Mohali (MPIC/IISER)). All stations except Mohali are part of NDACC. The technical problem regarding the Mainz data is resolved and data is available again.
- NDACC ZSL-DOAS stratospheric columns from 25 stations sample latitudes between 80°N (Eureka) and -75°S (Dome C).
- FTIR stratospheric columns up to March 2021, from 23 stations between 80°N (Eureka) and -78°S (Arrival Heights).
- PGN total column direct-sun data are available from 36 stations at the VDAF Automated Validation Server sample latitudes from 78°N to -46°S. Currently PGN is undergoing a version upgrade: PGN v1.7 is being replaced by v1.8, leading to a mix of versions at EVDC. For the current report, we therefore use data downloaded from the Automated Validation Server at 2022/05/03, when PGN was still fully at v1.7.
7.3.3 Tropospheric NO₂ column

7.3.3.1 Bias

The OFFL NO₂ tropospheric column values are compared with data from 27 NIDFORVAL MAX-DOAS stations. The median bias over all stations is -1.9 Pmolec/cm² (-34%). A summary of the bias and spread for all stations is shown in Figure 32 and Figure 33. The median bias for the subset of 8 MAX-DOAS stations in the VDAF-AVS (inspection 2022/05/30, 4863 co-locations) is -1.2 Pmolec/cm² (-26%) with a Pearson correlation of 0.76. For the PAL NO₂ tropospheric columns, we see a reduced bias of -0.7 Pmolec/cm² (-16%), and a Pearson correlation of 0.78 for 4352 co-locations. These results are within the mission requirement of a maximum bias of 50%.

With a station-to-station dispersion (NIDFORVAL, IP68/2 over all station medians) of 2.1 Pmolec/cm², a single bias number for all stations has limited meaning. Three regimes can be identified for comparisons, as discussed in Verhoelst et al. (2021): (1) low tropospheric NO₂ values (median tropospheric column <2 Pmolec/cm²), (2) polluted stations (from 3 to 14 Pmolec/cm²), and (3) extremely polluted stations (>15 Pmolec/cm²). The median bias in these regimes is about 0.14 Pmolec/cm² (15%), -1.9 Pmolec/cm² (-34.6%), and -10 Pmolec/cm² (-45.6%), respectively.

Figure 32: Box-and-whisker plots summarizing the bias and spread of the difference between SSP TROPOMI RPRO+OFFL and MAX-DOAS NO₂ tropospheric columns (left: absolute, right: relative). Black corresponds to data up to and including processor version 1.3. Green corresponds to processor version 1.4 and above, in operation since December 2020. Values between brackets in the labels denote the median tropospheric column at the station. The time frame is from May 2018 until end March 2022. Stations are ordered by median tropospheric column. The median difference is represented by a vertical solid line inside the box, which marks the 25 and 75% quantiles. The whiskers cover the 9-91% range of the differences. The shaded area represents the mission requirement 50% for the uncertainty.
**Figure 33**: Time series – from May 2018 until April 2022 – of the weekly averaged relative difference [%] between S5P RPRO+OFFL and MAX-DOAS NO\textsubscript{2} tropospheric column data. The black solid, dotted, dashed, and 2\textsuperscript{nd} dotted lines indicate processor switches from 1.2.2 to 1.3.0, to 1.4.0, to 2.2.0, and to 2.3.1, respectively, and the white dashed line the activation of the finer horizontal resolution. Stations are ordered by median tropospheric column, in brackets.

A test case study has been performed for ground-based MAX-DOAS stations providing low tropospheric profiles (with data up to end March 2022), which is the case for the BIRA-IASB, ChibaU, JAMSTEC, and UNAM stations. In this case, the TROPOMI averaging kernel can be used to smooth the MAX-DOAS profiles and remove the profile contribution from the comparison, using the following formula:

$$VCD_{\text{smoothed}} = AK_{\text{sat}} \times x_{\text{MAXDOAS}}$$

This analysis is illustrated on the 14 stations shown in **Figure 34** where the absolute and relative differences are presented for original comparisons (grey boxes) and after smoothing of the ground-based data (black boxes). The median bias then generally decreases absolutely (for 10 cases over 14 stations) by about 10 to 20% (see bottom of **Figure 34**). This reduction comes sometimes at the expense of a larger spread of the comparisons (e.g. Vallejo and Unam). The large increase of 40 and 60% for Fukue, Haldwani, and Cape Hedo can be attributed to the relative bias calculation from dividing with small tropospheric VCDs at very clean ground-based stations and the very small number of colocations (5) for the Haldwani station.
In summary, the tropospheric NO$_2$ bias depends on pollution level. On average, it is about -34%. It can be as high as -50% for extreme pollution and +11% in clean areas. Taking the sensitivities of the instruments into account by using the satellite averaging kernels with the low tropospheric MAX-DOAS profiles, the bias can be reduced by up to 20% absolutely.

7.3.3.2 Dispersion

The median dispersion in comparison to NIDFORVAL MAX-DOAS is 2.6 Pmole/cm$^2$. The median IP68/2 dispersion for the different pollution levels as defined in the previous section is (1) 0.66 Pmole/cm$^2$, (2) 2.6 Pmole/cm$^2$, and (3) 7.3 Pmole/cm$^2$, respectively. The uncertainty precision requirement of maximum 0.7 Pmole/cm$^2$ is only satisfied for the clean-station ensemble. A comparison of PAL and OFFL data shows similar dispersions.
It must be noted that MAX-DOAS uncertainty sources and comparison errors also contribute to the dispersion. Moreover, systematic errors (e.g., seasonal cycle) can contribute. A part of the systematic error component can be removed by calculating the dispersion (IP68/2) around the OLS regression line instead of the dispersion between S5P and MAX-DOAS data (adapted from Schneider et al., 2006). The residual dispersion IP68/2 is 0.9 Pmolec/cm² at Mohali, approximately 1.5 Pmolec/cm² at the different European VDAF-AVS stations, and 2.9 Pmolec/cm² at Xianghe. There is a reasonably good correlation between TROPOMI and MAX-DOAS tropospheric columns with the Pearson R varying between 0.6 (Bremen) and 0.84 (Xianghe) with a mean of 0.73.

7.3.3.3 Dependence on influence quantities

Two key influence quantities for observations of tropospheric NO₂ are aerosol optical depth (AOD) and satellite cloud (radiance) fraction (CRF). The first one is retrieved within the MAX-DOAS NO₂ analysis. When binning the MAX-DOAS comparison biases of each station by AOD from 0 to 2 in intervals of 0.5 and CRF from 0 to 0.5 in intervals of 0.05, a median bias increase towards larger bin values is found (Figure 35).

7.3.3.4 Seasonal and shorter term variability

Global zonal daily mean NO₂ tropospheric columns are shown in Figure 36. After the V01.04.00 change, a small column increase in the Northern hemisphere between 30° and 60°N is detectable. Overall, no trends or short term jumps in columns are visible.
Figure 36: SSP TROPOMI NO$_2$ tropospheric columns [$10^{15}$ molec/cm$^2$] as a function of day and latitude. The period is from May 2018 to May 2022. Grid box size in latitude direction is 0.5°. The grey vertical lines mark processor version changes, the black lines the beginning of each year.

Figure 37 presents the seasonal cycle of differences between S5P RPRO+OFFL and MAX-DOAS tropospheric NO$_2$ from the AVS. All comparison pairs are reported on a single year. It should be noted that the RPRO is version 01.02.02 for the period of May-October 2018. TROPOMI measures lower values than MAX-DOAS in late fall and winter, when tropospheric NO$_2$ reaches its largest abundance. Over the entire year, the 30-day rolling median relative difference is within the mission requirements for the bias.
Figure 37: Seasonal cycle (with data mapped to one generic year) of the difference between SSP RPRO+OFFL and MAX-DOAS NO2 tropospheric column data at six European stations. Difference (left) and relative difference (right). On the left column, the lowest 2.5% data is not shown for visibility. Data was obtained from the VDAF Automated Validation Server on 2022/05/25.
7.3.3.5 Geographical patterns

In general, geographical patterns or artefacts cannot be detected in OFFL v2.3.1 L2_NO2, as shown in the monthly means of January 2022 over central Europe (Figure 37) in comparison to 2020 (V1.3.2). But a substantial column increase in Northern Europe is observed for v2.3.1.

![Figure 37: S5P OFFL tropospheric NO2 over central Europe. The data is binned on a grid of 0.06° latitude and 0.03° longitude for January 2020 (V1.3.2) and 2022 (V2.3.1). A qa_value > 0.75 and a cloud fraction CF<0.6 are used to reduce the amount of data and exclude cloudy scenes. 1 PMC means 1 Pmolec/cm².](image)

**Figure 38**: S5P OFFL tropospheric NO2 over central Europe. The data is binned on a grid of 0.06° latitude and 0.03° longitude for January 2020 (V1.3.2) and 2022 (V2.3.1). A qa_value > 0.75 and a cloud fraction CF<0.6 are used to reduce the amount of data and exclude cloudy scenes. 1 PMC means 1 Pmolec/cm².

7.3.3.6 Processor changes from version 1.3 to versions 1.4.0 and 2.3.1

Changes related to the version change v1.4 are clearly detectable for some stations (Figure 32). To quantify this effect, we calculate the median relative differences and its dispersions for all the measurement pairs for the January-February-March period of 2022 (v2.3.1), 2021 (v1.4), 2020 (v1.3), and 2019 (mix of v1.2 and v1.3). The results are given in Table 5. The v1.4 change leads to an absolute reduction of about 10% that started in December 2020. Version 2.3.1 leads to smaller biases with a reduction of another 9%. As V2.2 and V2.3.1 share the same algorithm, no changes in bias are expected.

**Table 5**: Median relative difference and dispersion between S5P RPRO+OFFL and MAX-DOAS NO2 tropospheric column data, for the January-February-March periods of 2019 (mix of v1.2 and v1.3), 2020 (v1.3), 2021 (v1.4), and 2022 (v2.3.1).

<table>
<thead>
<tr>
<th>Date</th>
<th>JFM 2019</th>
<th>JFM 2020</th>
<th>JFM 2021</th>
<th>JFM 2022</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version</td>
<td>1.2/1.3</td>
<td>1.3</td>
<td>1.4</td>
<td>2.3.1</td>
</tr>
<tr>
<td>Collocation count</td>
<td>793</td>
<td>818</td>
<td>692</td>
<td>320</td>
</tr>
<tr>
<td>Median bias [%]</td>
<td>-38.5</td>
<td>-37.2</td>
<td>-27.1</td>
<td>-18</td>
</tr>
<tr>
<td>Dispersion (0.5 IQ68) [Pmolec/cm²]</td>
<td>5.3</td>
<td>4.1</td>
<td>3.5</td>
<td>2.5</td>
</tr>
</tbody>
</table>

A similar comparison of the July-December periods 2019, 2020 and 2021 shows a combined absolute bias decrease of roughly 4% at 5 stations, but for less collocations (#360).

Comparisons of S5P NO₂ v1.2 and v1.3 with the QA4ECV OMI NO₂ retrieval product reveal substantial seasonal differences in the tropospheric columns, with TROPOMI systematically lower than OMI. The main difference between the two retrievals is the cloud pressure retrieval. OMI uses the O₂-O₂ absorption feature and TROPOMI the O2-A absorption and FRESCO-S cloud retrieval approach. In particular, the FRESCO-S version implemented for TROPOMI in v1.0 to v1.3 is known to overestimate the cloud pressure, leading to a high bias in the air-mass factors and a low bias in the tropospheric columns.
On 2 December (NRTI) and 29 November (OFFL) 2020, the SSP NO\textsubscript{2} product was upgraded from v1.3.2 to v1.4.0. In versions 1.4 and beyond, the FRESCO-S cloud retrieval was upgraded to include the weaker absorbing wavelengths in the O-2A absorption band. This results in systematically lower cloud pressures, lower air-mass factors, and higher tropospheric columns. Comparisons of SSP NO\textsubscript{2} v1.4 and v2.2.0 (van Geffen et al., 2022, figures 13 and 14) with the QA4ECV OMI NO\textsubscript{2} show a much-improved consistency between the two retrievals. The large difference over China in v1.3 and the corresponding improvement of v1.4 is demonstrated in Figure 39. For v1.4, the comparison with OMI shows much better agreement of the tropospheric NO\textsubscript{2} columns.

![Figure 39: TROPOMI-OMI tropospheric NO\textsubscript{2} monthly scatter plots for Eastern China [110E-124E;21N-43N]. Shown are the tropospheric NO\textsubscript{2} column (VCD\textsubscript{trop}) of TROPOMI (y-axis) versus OMI (x-axis) for December and January 2019/2020 (top; v1.3) and 2020/2021 (bottom; v1.4). TROPOMI data for pixels with qa\_value > 0.75 and OMI/QA4ECV data with cloud radiance fraction < 0.5 (i.e. "clear-sky pixels") have been gridded to a common grid of 0.8°lon x 0.4°lat and are averaged over calendar months.](image)

### 7.3.3.7 Other features

Ordinary linear regression (OLS) of SSP vs MAX-DOAS yields fairly good correlation coefficients (0.59-0.85, see before), but low slopes. When \(SSP = a \times MXD + b\), \(a\) varies between 0.3 (Bremen) to 0.6 (Athens). It is known, however, that this approach is only correct in the limit that all random errors are in SSP data. For the MAX-DOAS vs SSP OLS (i.e., assuming the opposite limit that all random errors are in MAX-DOAS), still with \(SSP = a \times MXD + b\), one obtains slopes closer to unity: \(a\) varies then between 0.74 (Mainz) and 0.99 (Uccle).
7.3.4 Stratospheric NO$_2$ column

7.3.4.1 Bias

NRTI stratospheric NO$_2$ column values are generally lower than the photochemical corrected ground-based ZSL-DOAS values by approximately -0.17 Pmolec/cm$^2$ (-6.1%), with a station-to-station scatter of the mean bias of similar magnitude (Figure 40). The validation results are based on comparisons at 25 NDACC ZSL-DOAS stations using a photochemical correction and sampling the latitude range from 80°N (Eureka) to -75°S (Dome C).

For the subset of 13 ZSL-DOAS stations contributing data to the VDAF-AVS (10923 co-locations, inspection date 2020/05/25), the OFFL stratospheric NO$_2$ column median bias is -4.6% (-0.14 Pmolec/cm$^2$) with a station-to-station 1σ scatter of 7.8% and a high Pearson correlation of 0.91. Both stratospheric column results are within the mission requirement of 10% maximum bias (equivalent to 0.2-0.4 Pmolec/cm$^2$, depending on latitude and season). The bias decreased with version 2.2.0 to -3%.

Figure 40: Pole to pole bias and spread [Pmolec/cm$^2$] of S5P TROPOMI NRTI and NDACC ZSL-DOAS NO2 stratospheric columns (SAOZ data in black, other ZSL-DOAS in blue, and mountain-top PGN in red). The network-wide mean bias and its formal uncertainty are -0.17 and 0.06 Pmolec/cm$^2$, respectively. Stations appearing twice have had their data processed both in NRT mode (LATMOS_RT) and with the LATMOS_v3 processor. Both are shown in order to illustrate the consistency. This graph includes results for processor versions 1.1.0 and later. The median difference is represented by a vertical solid line inside the box, which marks the 25 and 75% quantiles and the whiskers the 9-91% range. The red shaded area represents the mission requirement of 0.5 Pmolec/cm$^2$ for the uncertainty. Station name and latitude of the station are added on the left. The time frame is May 2018 to February 2022.
Figure 41: Time series – from August 2018 until May 2022 – of S5P NRTI L2_NO2 and NDACC ZSL-DOAS NO2 stratospheric column differences, weekly averaged [Pmolec/cm²]. The black solid, dotted, dashed, and 2nd dotted lines indicate, respectively, the processor switches from 1.2.2 to 1.3.0, to 1.4.0, to 2.2.0, and to 2.3.1, and the white dashed line indicates the activation of the finer horizontal resolution. For stations appearing twice, ground-based data were processed both in NRT mode (LATMOS_RT) and with the LATMOS_v3 processor to verify the consistency.

The median bias between S5P and FTIR NO2 stratospheric columns is +4.7% for the combined 24 stations, but with larger biases at high latitude and tropical stations (9-14%, see Figure 40). The station-to-station 1σ scatter is 5.9%. The reason for the discrepancy between FTIR and ZSL-DOAS is on-going work.

Figure 42: Median bias at each NDACC FTIR station as a function of latitude, calculated as the median of the percentage difference between S5P L2_NO2 and the FTIR NO2 stratospheric column measurement. The grey bars are the scaled MAD (see Sect. 6.2.1), and the coloured bars are the ±2σ error on the bias.
7.3.4.2 Dispersion

From ZSL-DOAS comparisons, the ±1σ dispersion of the difference between stratospheric column and reference data around their median value rarely exceeds 0.3 Pmolec/cm² at stations without tropospheric pollution (cf. the error bars in Figure 40). When combining random errors in the satellite and reference measurements with irreducible collocation mismatch effects, it can be concluded that the random uncertainty on the S5P stratospheric column measurements falls within mission requirements of maximum 0.5 Pmolec/cm². The mean Pearson-R is 0.96±0.1.

Similar conclusions are reached from FTIR comparisons. The scaled MAD (equivalent to 1σ dispersion) of the differences is 0.29 Pmolec/cm² for all data together, with a Pearson correlation coefficient of 0.93. At individual stations, it never exceeds 0.5 Pmolec/cm², except at Toronto (0.6) and Paramaribo (0.8). The robust correlation coefficient is 0.97, as shown in the scatter plot (Figure 43, “Correlation MAD”).

![Figure 43: Scatter plot of co-located S5P TROPOMI and NDACC FTIR stratospheric NO₂ column data.](image)

7.3.4.3 Dependence on influence quantities

The evaluation of potential dependences of the S5P stratospheric column on Solar Zenith Angle (SZA), cloud fraction (CF) and surface albedo of the SSP measurement does not reveal bias variations much larger than 0.4 Pmolec/cm² over the range of the influence quantities.
7.3.4.4 Seasonal cycle and shorter term variability

TROPOMI and ground-based ZSL-DOAS instruments both capture the short-term variabilities (at daily and monthly scales) of the NO$_2$ stratospheric column, as illustrated for the NDACC station of Kerguelen Island in Figure 45. The ground-based SAOZ data acquired at twilight were adjusted to account for the photochemical diurnal variation between twilight and the early afternoon S5P overpass time.

Figure 44: Difference between S5P L2 NO$_2$ NRTI and ground-based SAOZ stratospheric NO$_2$ columns as a function of the satellite solar zenith angle (SZA), satellite cloud fraction, and satellite surface albedo. Mean and standard deviation calculated over bin widths of 10° degrees in SZA, 0.1 in CF, and 0.1 in surface albedo (solid black line and grey bars). Co-locations cover the period from May 2018 to August 2021.

Figure 45: Time series of S5P NRTI L2 NO$_2$ stratospheric NO$_2$ column data (blue dots) co-located with ground-based SAOZ twilight measurements (red dots) at sunset performed by LATMOS at the NDACC southern mid-latitude station of Kerguelen Island. In the upper plot, the photochemical correction is deactivated to offset the two time-series and to better see the day-to-day variability. The time frame is July 2018 to May 2022.
TROPOMI stratospheric NO$_2$ and Pandora total NO$_2$ at three mountain stations plotted follow the same seasonal cycle (Figure 46). It must be noted that the PGN Pandora data is more scattered. The 30-day rolling median of the relative difference is within the bias requirements, except for the months 2019 January-March at Mauna Loa and 2020 March-April at Izaña. But in both cases this is due to Pandora measurements not following a regular seasonal cycle.

Figure 46: Time series of the S5P RPRO-OFFL L2_NO2 stratospheric NO$_2$ column and Pandora total column at 3 mountain stations. Left row: S5P and PGN Pandora together; centre: their difference; right: their percent relative difference. Plain lines represent the 30-day rolling mean or (for relative difference) the 30-day rolling median. Black full vertical line, dotted line, dashed and dot-dashed line: OFFL processor version change 1.2 to 1.3 on 2019/03/20, 1.3 to 1.4 on 2020/11/29, 1.4 to 2.2 on 2021/07/01 and 2.2 to 2.3 on 2021/11/14, respectively. Green dashed line: pixel size switch at 2019/08/06. Data covers the period from June 2018 to February 2022. It is obtained from the validation server on 2022/02/28.

7.3.4.5 Geographical patterns

None to report.
7.3.4.6 Processor version changes and switch to smaller ground pixel resolution

The effect of (1) upgrades of the NL-L2 processor to v1.3.0 (2019-07-20), v1.4.0 (2020-12-02, OFFL) v2.2.0 (2021-07-01) and to v2.3.1 (2021-11-14), and (2) the change in TROPOMI ground pixel size on 6 August 2019 on the S5P stratospheric NO2 column data was investigated by comparing the S5P and ground-based time series at the NDACC ZSL-DOAS stations (Figure 41) and 3 PGN Pandora mountain stations (Figure 46). The difference between the S5P and ground-based data does not show any impact from the processor and pixel size changes. The significant processor upgrade from v1.3 to v1.4 in December 2020 affects mainly the tropospheric part of the column. But also the stratospheric column is slightly affected by the upgrades. A comparison of 3 similar time spans (July/December 2019-2021) was performed and a decrease of the negative bias (-6%) in 2019 and 2020 to -4.6% in 2021 (9 stations, #1059 measurements) was detected. The v2.2 upgrade using L1B V2 input data lead to a small increase in stratospheric columns (van Geffen et al., 2022), explaining partly the bias decrease.

7.3.4.7 Other features

None to report.

7.3.5 Total NO2 column

7.3.5.1 Bias

Based on measurements from 36 Pandora stations between 78.9°N and -45.8°S available at VDAF-AVS (Figure 47), the median bias (20502 co-locations, checked 2022/02/25) is -6.7% (-0.9 Pmolec/cm²) with a station-to-station scatter of 18% and a Pearson correlation coefficient of 0.64. The results are within the 30% accuracy requirement, which is the average of the tropospheric and stratospheric bias maxima. Sorting the stations by pollution level (6 Pmolec/cm²), the bias is +4% for the 20 low polluted stations and -21% for 16 high polluted stations.
Figure 47: Box-and-whisker plots summarizing the bias and spread [Pmolec/cm²] (left) and relative bias and spread [%] (right) between S5P TROPOMI RPRO+OFFL (processor version 1.2.2 up to 1.3.2 in black, version 1.4.0 and above in green) and PGN Pandora NO₂ total column data. Conventions of the boxplots are identical to Fig. 21. Stations are ordered by median total column. The time frame is May 2018 to April 2022.

We highlight three different comparison cases here. At Alice Springs (Australia), where the total NO₂ column values are mostly between 2-4 Pmolec/cm², a small positive mean bias of 0.2 Pmolec/cm² is seen (:+8% median bias). A wider distribution of NO₂ values (2-30 Pmolec/cm²) is found at Bronx (New York, United States) where the bias is -3 Pmolec/cm² (-13%). Finally, at Sapienza (Rome, Italy) where column values can reach up to 40 Pmolec/cm², the bias is -4 Pmolec/cm² (-29%), probably due to locally enhanced NO₂. The bias has been reduced in the months after the V1.4.0 change.

7.3.5.2 Dispersion

The dispersion of the S5P and PGN Pandora differences depends strongly on the station. Small dispersions (IP68/2) are observed at Ny-Ålesund, Alice Springs, Mauna Loa, Pilar, Izaña, Comodoro Rivadavia (0.4-0.6 Pmolec/cm²) that are within the mission precision requirement, and higher values elsewhere (e.g., 3-5 Pmolec/cm² at New York Bronx, Sapienza Rome, New York City College, Unam). The per-station median of the IP68/2 is 1.3 Pmolec/cm². The mean Pearson-R is 0.64 and varies from relatively low (e.g., 0.39 at Charles City) to high (0.87 at Wakkerstroom).

7.3.5.3 Dependence on influence quantities

None to report.
7.3.5.4  **Short term variability**

None to report.

7.3.5.5  **Geographical patterns**

None to report.

7.3.5.6  **Processor version changes and switch to smaller ground pixel resolution**

Three major processor version changes (1.2 to 1.3 to 1.4, and to 2.2) occurred for OFFL L2_NO2 on 20 March 2019, 2 December 2020, and 1 July 2021, respectively. On 6 August 2019, there was a change in the TROPOMI ground pixel size. All changes were investigated by having a close look at the S5P and PGN Pandora time series at individual stations. Figure 48 shows that the bias and scatter of the difference are not affected by the pixel size change. The change to processor version 1.4.0 appears to lead to more positive values in the differences (and thus to a less pronounced negative bias at polluted stations), as expected from the improvements in the cloud parameters.
To quantify this effect, we calculate the median relative difference and the difference dispersion for all measurement pairs for the January-February-March period of 2021 (v1.4), 2020 (v1.3) and 2019 (mix of v1.2 and v1.3). Results are given in the Table 6. Bias and dispersion appear lower in 2021, but note the lower number of comparison pairs here, which could be due to less stations contributing data in this period, or a higher qa_value filter impact.

The processor changes v2.2.0 and v2.3.1 do not introduce any discontinuities. The check for the January to March periods shows an absolute decrease of the median bias to -3.8% in 2022 with similar dispersion (1.8 Pmolec/cm²) for 19 out of 36 stations (~5000 measurements per year).

Table 6 - Median relative difference and its dispersions between S5P RPRO+OFFL and Pandora NO2 total column data, for the January-February-March period of 2021 (v1.4), 2020 (v1.3) and 2019 (mix of v1.2 and v1.3). A qa_value > 75 % is applied.

<table>
<thead>
<tr>
<th>Date</th>
<th>JFM 2019</th>
<th>JFM 2020</th>
<th>JFM 2021</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version</td>
<td>1.2/1.3</td>
<td>1.3</td>
<td>1.4</td>
</tr>
<tr>
<td>Collocation count</td>
<td>859</td>
<td>978</td>
<td>591</td>
</tr>
<tr>
<td>Median [%]</td>
<td>-5.9</td>
<td>-6.5</td>
<td>-2.3</td>
</tr>
<tr>
<td>Dispersion (0.5 IQ68) [Pmolec/cm²]</td>
<td>2.6</td>
<td>2.3</td>
<td>1.8</td>
</tr>
</tbody>
</table>

7.3.5.7 Other features

None to report.
7.4   Equivalence of L2_NO2 NRTI and OFFL products

This section shows evidence that the L2_NO2 NRTI and OFFL products do not differ significantly and that their respective validations yield similar conclusions. We show the differences between the two datasets for the three different products (stratospheric, tropospheric, and total column).

7.4.1   Tropospheric NO2 Column NRTI vs OFFL

To demonstrate the closeness of the NRTI and OFFL L2_NO2 products at MAX-DOAS stations, L2_NO2 NRTI (processor version 01.00.02 to 02.03.01) and L2_NO2 OFFL (RPRO processor version 01.02.02 + OFFL processor version 01.02.00 to 02.03.01), each co-located with MAX-DOAS, were obtained from the validation server, and the subset of pixels, common to both NRTI and OFFL, was determined. Statistical results for Bremen and Mainz are summarized in Table 7. Similar conclusions on the closeness of NRTI and OFFL can also be drawn for the other stations. V2.3.1 starting in the middle of November is not used here.

Table 7 - Statistics on the comparison of the common subset of L2_NO2 NRTI, L2_NO2 RPRO+OFFL and co-located MAX-DOAS, for the stations Bremen and Mainz (*: unit of Pmolec/cm2).

<table>
<thead>
<tr>
<th></th>
<th>NRTI-OFFL</th>
<th>NRTI-MXD</th>
<th>OFFL-MXD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bremen: 594 common co-locations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean(diff) ±sem*</td>
<td>0.05</td>
<td>-2.12±0.14</td>
<td>-2.17±0.14</td>
</tr>
<tr>
<td>Median(diff)*</td>
<td>0.06</td>
<td>-1.31</td>
<td>-1.34</td>
</tr>
<tr>
<td>Std(diff)*</td>
<td>0.4</td>
<td>3.5</td>
<td>3.5</td>
</tr>
<tr>
<td>1/2 IP68(diff)*</td>
<td>0.3</td>
<td>2.3</td>
<td>2.3</td>
</tr>
<tr>
<td>Pearson R</td>
<td>0.98</td>
<td>0.59</td>
<td>0.59</td>
</tr>
<tr>
<td>Slope</td>
<td>0.99</td>
<td>0.30</td>
<td>0.29</td>
</tr>
<tr>
<td>Mainz: 447 common co-locations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean(diff) ±sem*</td>
<td>0.08</td>
<td>-3.11±0.17</td>
<td>-3.19±0.16</td>
</tr>
<tr>
<td>Median(diff)*</td>
<td>0.08</td>
<td>-2.29</td>
<td>-2.41</td>
</tr>
<tr>
<td>Std(diff)*</td>
<td>0.5</td>
<td>3.5</td>
<td>3.4</td>
</tr>
<tr>
<td>1/2 IP68(diff)*</td>
<td>0.2</td>
<td>3.2</td>
<td>3.2</td>
</tr>
<tr>
<td>Pearson R</td>
<td>0.99</td>
<td>0.78</td>
<td>0.79</td>
</tr>
<tr>
<td>Slope</td>
<td>0.99</td>
<td>0.48</td>
<td>0.49</td>
</tr>
</tbody>
</table>

The mean difference between NRTI and OFFL is of the same order or smaller as the standard error on the mean difference of NRTI-MAX-DOAS and OFFL-MAX-DOAS. Therefore, the bias difference between NRTI and OFFL is statistically not significant. Also, the difference dispersion between NRTI and OFFL is small compared to the difference dispersion between either NRTI or OFFL on the one hand and MAX-DOAS on the other hand. The good match between NRTI and OFFL is also demonstrated by the high Pearson R values and the near unity slope of the linear regression.

Daily tropospheric columns from NRTI and OFFL V2.3.1 data streams are binned to 0.5° grid cells and differentiated (Figure 49) to elucidate the spatial distribution of differences. The Pearson correlation coefficient is high (0.98) and the global relative mean difference is very small with 0.6 %. Differences above ±1 Pmolec/cm² are rarely found.
7.4.2 Stratospheric NO\textsubscript{2} Column NRTI vs OFFL

The similarity of the two products can be investigated by comparing the processing of a randomly chosen orbit. Figure 50 shows this approach for August 27, 2021. It reveals differences mostly below the mission requirement on precision (0.5 Pmolec/cm\textsuperscript{2}). The RMSD is 0.09 Pmolec/cm\textsuperscript{2}, with values up to 0.5 Pmolec/cm\textsuperscript{2}. Some features are due to a different stratosphere/troposphere separation (e.g. over Greenland, positive difference in stratosphere, negative in troposphere, Figure 50). Since these differences, representing up to 20% of the stratospheric column, do exceed the mission requirement on the bias (10%), and because a much more comprehensive orbit-by-orbit analysis is needed to ensure differences remain reasonable under all conditions, the full validation analysis as performed for the NRTI product was repeated on the OFFL product.

A comparison of stratospheric NRTI and OFFL pole-to-pole validation graphs as shown in Figure 51 illustrates that in direct validation studies the NRTI performs very similarly to OFFL with a median bias of less than -0.2 Pmolec/cm\textsuperscript{2} at sunset.
7.4.3 Total NO₂ Column NRTI vs OFFL

The global relative total column difference is in the range of ±0.1 Pmolec/cm² (2022/05/15), with NRTI being slightly higher. The Pearson correlation coefficient between both data sets is 0.99. The comparison as previously analysed for the stratospheric and tropospheric columns reveals that most of the features seen in the sub-columns compensate each other in the total columns, indicating a different troposphere-stratosphere separation (Figure 52). A small disparity within an orbit between East and West in the latitude band of 40°N/S are found in the range of ±0.1 Pmolec/cm².

Figure 51: Meridian dependence of the mean (the circular markers) and dispersion (±1σ error bars) of the differences between S5P TROPOMI L2 NO₂ (NRTI in the left panel, OFFL in the right) stratospheric column data and ZSL-DOAS reference data, represented at individual stations from the Antarctic to the Arctic. The values in the legend correspond to the median and its formal uncertainty for all mean (per station) differences. The figure covers the period from the start of phase E2 to August 2021.

Figure 52: Difference between S5P NRTI and OFFL V2.3.1 daily total NO₂ column data (binned to 0.5°x0.5° resolution) for 2022/05/15. The difference is only calculated for columns above 0.5 Pmolec/cm², cloud fractions below 0.6 and a qa_value above 0.75.
7.5 Internal consistency of the NO$_2$ validation results

This section focuses on the internal consistency checks of the NO$_2$ validation results. This relies on analysis of potential source of inconsistencies (such as impact of different NO$_2$ cross-sections) or synergistic analysis of several instruments types.

7.5.1 NO$_2$ absorption cross-sections

A potential source of inconsistencies between the different data products lies in the NO$_2$ absorption cross sections that are used in the DOAS retrieval of the slant column density (SCD). An overview of the different NO$_2$ cross sections choices made for each instrument is provided in Table 8, reproduced from Verhoelst et al. (2021). For a detailed discussion we refer to this work. The main conclusions are:

- A small (few percent) seasonal cycle in the stratospheric column comparisons can be expected, due to the seasonal variation in stratospheric temperature not being accounted for in the ZSL-DOAS data processing.
- PGN columns may either overestimate by up to 10% when the column is mostly stratospheric or underestimate by a similar order of magnitude when large tropospheric amounts are present, due to the use of a fixed effective temperature of 254.4 K.
- The MAX-DOAS data may be biased in either direction by a few percent when tropospheric and/or stratospheric temperatures differ strongly from the 298 K and 220 K default temperatures.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>reference</th>
<th>temperature</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>S5p TROPOMI</td>
<td>Vandeule et al. (1998)</td>
<td>220K</td>
<td>With temperature correction in AMF (Zara et al., 2017)</td>
</tr>
<tr>
<td>ZSL-DOAS</td>
<td>Vandeule et al. (1998)</td>
<td>220K</td>
<td>NIWA instruments</td>
</tr>
<tr>
<td>ZSL-DOAS</td>
<td>Harder et al. (1997)</td>
<td>227K</td>
<td></td>
</tr>
<tr>
<td>MAX-DOAS</td>
<td>Vandeule et al. (1996)</td>
<td>298K</td>
<td>tropospheric retrieval only</td>
</tr>
<tr>
<td>PGN</td>
<td>Vandeule et al. (1998)</td>
<td>254.4K</td>
<td>PGN processor v1.7</td>
</tr>
</tbody>
</table>

7.5.2 ZSL-DOAS and PGN with low pollution level (mountain-top, arctic)

Three of the PGN direct-sun instruments are located near the summit of a volcanic peak: Altzomoni (3985m a.m.s.l) in the State of Mexico, Izaña (2360m a.m.s.l.) on Mount Teide on the island of Tenerife, and Mauna Loa (4169m a.m.s.l.) on the island of Hawaii. At these high-altitude stations, the total column measured by the ground-based direct-sun instrument misses most of the tropospheric (potentially polluted) part and as such becomes representative of the TROPOMI stratospheric column (with a minor free tropospheric column part). These stations have been used in Verhoelst (2021) for the stratospheric comparison. As for the zenith-sky data, a minor negative median difference (TROPOMI-GB) of the order of -0.2 Pmolec/cm$^2$ was detected.

As mentioned, the PGN data are processed using cross sections at a single temperature, leading to columns which are about 10% larger than if they had been processed with cross sections at 220 K. Future processing of the PGN data will address this, and it is expected that this will mostly remove the apparent negative bias for TROPOMI but lead to a slight inconsistency with the ZSL-DOAS results, even more so as mountain-top PGN is also sensitive to free troposphere NO$_2$. This needs to be confirmed with a future PGN release.
At arctic stations (Eureka, Ny-Ålesund), where the tropospheric contribution to the column is expected to be small, both ZSL-DOAS and PGN instruments are located. All instruments follow the temporal evolution of S5P NO$_2$ rather well. However, there is a clear negative bias (TROPOMI lower than PGN) of about -0.8 Pmolec/cm$^2$ or -15% for the two PGN instruments, and a much smaller to no negative bias for the ZSL-DOAS instruments (no bias at Eureka, -0.4 Pmolec/cm$^2$ at Ny-Ålesund when considering the same time period). Also here, it is expected that the upcoming PGN release will reduce this discrepancy.

### 7.5.3 Stations with multiple instruments (different geometries)

There are a number of stations that have several instruments, covering different viewing geometries, and thus allowing an investigation of the internal consistency of the validation results. Direct-sun and MAX-DOAS instruments measure total and tropospheric NO$_2$ columns, respectively. By subtracting the TROPOMI stratospheric VCD from the direct-sun total columns, an estimation of the tropospheric NO$_2$ can be obtained and compared to MAX-DOAS results (Pinardi et al., 2020). Past comparisons (e.g., Figure 4 of Pinardi et al., 2020) point to a good consistency, with high correlations and biases of 10 to 15%. Current stations with these two types of instruments are Athens, Thessaloniki, Xianghe, Uccle, Yokosuka and Unam.

Preliminary comparisons have been done, but there is insufficient data for consolidated conclusions so far, due to impact of the seasonality, short overlapping periods and the need to consider differences in horizontal and vertical NO$_2$ representatively.

### 7.5.4 Consistency between MAX-DOAS and PGN network results

Another way to explore the consistency of the validation results is to compare the results at the network level, by comparing TROPOMI tropospheric columns to the MAX-DOAS values on one hand (cf. Section 6.3.3) and on the other hand to the calculated PGNtropo values (PGN total columns minus collocated TROPOMI stratospheric columns). Figure 52 presents the mosaic plots of both tropospheric comparisons. It can be seen that both networks present a positive bias for clean stations (i.e. tropospheric VCD <1.8 Pmolec/cm$^2$ for MAX-DOAS stations), and a negative bias for more polluted stations. For MAX-DOAS comparisons, the positive-to-negative bias crossing appears to occur between 1.2 Pmolec/cm$^2$ (fukue_J) and 1.8 Pmolec/cm$^2$ (Phimai), while for PGNtrop it occurs between 1.9 Pmolec/cm$^2$ (Canberra) and 2.6 Pmolec/cm$^2$ (Boulder) with station-dependent negative and positive biases in-between. The MAX-DOAS network is heavily biased towards stations with significant tropospheric column, with a lack of stations in the 2 to 3 Pmolec/cm$^2$ range compared to the PGN network.

Some stations hosting both a MAX-DOAS and a Pandora instrument present some apparent inconsistencies. For instance, the Unam PGNtrop median value (15 Pmolec/cm$^2$) is quite below the MAX-DOAS one (between 19.9 and 21.1 Pmolec/cm$^2$, depending on the viewing directions), and the Athens PGNtrop median value (6.9 Pmolec/cm$^2$) is higher than the MAXDOAS one (3.7 Pmolec/cm$^2$). In the latter case however the 2 instruments are not located at the same station in Athens, with the MAXDOAS on top of a hill, missing the urban boundary layer contribution. As discussed in the previous subsection, more analysis should be done (keeping only common time periods, exploring the line of sight influence and the impact of the clouds, on the comparisons, in addition to comparing the GB datasets themselves before any TROPOMI collocation.
Figure 53: Time series – from May 2018 until April 2022 – of S5P RPRO+OFFL and a) MAXDOAS and b) PGNtropo NO$_2$ tropospheric column differences, weekly averaged [%]. The black solid, dotted, dashed, and 2nd dotted lines indicate, respectively, the processor switches from 1.2.2 to 1.3.0, 1.4.0, 2.2.0 and 2.3.1, and the white dashed line the activation of the finer horizontal resolution. Stations are ordered by median tropospheric columns (see numbers in brackets).
8 Validation Results: L2_HCHO

8.1 L2_HCHO products and requirements

This section reports on the validation of the following geophysical variables of the S5P TROPOMI L2_HCHO product identified in Table 1: the HCHO total column. Validation results are discussed with respect to the product quality targets outlined in Table 3. As the NRTI and OFFL processors are producing very similar data products, mainly the validation of the L2_HCHO OFFL product is reported hereafter. Subsection 8.4 shows evidence that NRTI and OFFL data do not differ significantly and that their respective validiations yield similar conclusions.

The operational (E2) phase for the S5P TROPOMI mission starts with orbit #02818 on April 30, 2018. The L2_HCHO reprocessed data product version 01.01.05 covers the period from 30 April to November 2018. L2_HCHO NRTI and OFFL product version 02.02.01 was released beginning of July 2021. The latest upgrade to version 02.03.00 (with improved background correction), issued in March 2022, will be assessed in a next update of this report.

8.2 Validation approach

8.2.1 Ground-based monitoring networks

S5P L2_HCHO data are routinely validated through comparisons with respect to ground-based measurements acquired by NDACC FTIR and MAX-DOAS UV-visible instruments performing network operation in the framework of NDACC. For S5P validation purposes those measurements are collected either automatically through EVDC or manually through SSPVT AO projects with faster data delivery (e.g., CESAR AO ID 28596 and NIDFORVAL AO ID 208607).

8.2.1.1 Fourier Transform Infrared Spectrometers

TROPOMI L2_HCHO formaldehyde column data are compared to reference measurements acquired at NDACC FTIR stations. FTIR measurements have a median systematic uncertainty of 13% and a median random uncertainty of 0.3 Pmolec/cm² (Vigouroux et al., 2018). The vertical sensitivity of FTIR is similar to that of S5P HCHO, with lower sensitivity close to the surface.

The comparison methodology is described in Vigouroux et al. (2020). Here we only give a brief outline:

- SSP pixels are selected within 20 km of the FTIR station (about 30-40 pixels). Only pixels with a qa_value > 0.5 are used. A collocation pair is only kept if at least 10 pixels can be averaged.
- The time coincidence criterion is set to ±3 hours of the satellite overpass time.
- The following data manipulations are performed: (i) The FTIR a priori profile is substituted with the TROPOMI L2_HCHO one to get a corrected FTIR profile. (ii) The corrected profile is smoothed with the TROPOMI averaging kernel (Rodgers and Connor, 2003). (iii) Scaling is applied to take into account altitude differences between pixel level and station altitude. (iv) Both the individual manipulated FTIR columns and the individual SSP manipulated pixel columns are then averaged.
- The relative median bias at a single station is estimated by the median relative difference: Med[|SAT-REF|/REF]. Absolute-scale dispersion is estimated by the scaled median absolute deviation from the median (MAD): 1.4826*MEDIAN[ABS(DIFF-MEDIAN(DIFF))]. The scaling factor of 1.4826 ensures that for a normal distribution, the MAD is equal to the standard deviation.
8.2.1.2 **MAX-DOAS UV-Visible Spectrometers**

TROPOMI L2_HCHO formaldehyde column data are routinely compared to reference measurements acquired by MAX-DOAS UV-Visible spectrometers. MAX-DOAS HCHO column data have a maximum bias 20% with a precision better than 30%. The MAX-DOAS vertical sensitivity differs from the S5P HCHO sensitivity. While MAX-DOAS has a higher sensitivity close to the surface and a lower sensitivity at higher altitudes, the reverse is true for S5P HCHO. Currently two channels are used to acquire MAX-DOAS data and perform the comparisons, each with their own comparison methodology:

- **VDAF Automated Validation server**: The S5p pixels are kept for $qa_value \geq 0.5$. It covers the MAX-DOAS measurement location, and is within ±0.5 h of the MAX-DOAS measurement. All MAX-DOAS measurements within ±0.5 h of the satellite overpass are averaged. No a priori substitution or averaging kernel is applied.
- **NIDFORVAL AO project**: S5P pixels with $qa_value \geq 0.5$ are kept. The average of S5P pixels within 20 km radius is compared with the average of MAX-DOAS measurements within ±3 h of satellite overpass. Note that these are the same co-location criteria as for the FTIR comparisons. For stations that also deliver an averaging kernel, a priori substitution followed by averaging kernel smoothing (Rodgers and Connor, 2003) is optionally applied. Relative bias and absolute-scale dispersion are calculated as for the validation based on FTIR data.

8.2.1.3 **Mutual consistency of the FTIR and MAX-DOAS ground-based data**

The Xianghe station (China, 39.75° N, 116.96°E) is one of the few stations where both FTIR and MAX-DOAS instruments are measuring in parallel since more than a year. In addition, there are also direct-sun measurements from the BIRA MAXDOAS and a second MAXDOAS instrument from USTC, and there are also plans to install a Pandora. It is thus an excellent candidate to test the consistency of the two techniques in a polluted station. Summary results (Figure 54) show very consistent results between direct-sun and FTIR data, a bit more spread for the direct-sun vs MAX-DOAS data (with larger MAX-DOAS values in winter time) and good regression but with large negative intercept for the FTIR vs MAX-DOAS original columns.

When taking into account the MAXDOAS and FTIR own sensitivities and using the Rodgers and Connor (2003) methodology with a priori substitution and smoothing, the initial MAX-DOAS-FTIR 27% bias is reduced to 15%, with a regression analysis showing a reduced slope, but also a much reduced intercept value between the two instruments (from $y = 0.97x - 2.1e15$ to $y = 0.89x - 0.2e15$). More information can be found in Pinardi et al. (MAXDOAS workshop meeting 2021).
Figure 54: a) Time-series of the MAX-DOAS, FTIR and direct-sun HCHO dataset from BIRA in Xianghe; Scatter-plot between the b) direct-sun and the MAX-DOAS, c) direct-sun and FTIR and d) FTIR and MAX-DOAS for the raw comparisons and when taking into account a-priori profile substitution and smoothing.

8.2.2 Satellites

TROPOMI L2_HCHO formaldehyde column data are also compared to similar data from the Metop-A and B GOME-2 data (version GDP 4.8) and the EOS-Aura Ozone Monitoring Instrument (OMI) using the QA4ECV L2 product (http://doi.org/10.18758/71021031).

8.2.3 Field campaigns and modelling support

Nothing to report.

8.3 Validation of L2_ HCHO

8.3.1 Recommendations for data usage

In order to avoid misinterpretation of the data quality, only those TROPOMI pixels associated with a qa_value > 0.5 (no error flag, cloud radiance fraction at 340 nm < 0.5, SZA equal to or below 70°, surface albedo smaller than or equal to 0.2, no snow/ice warning, air mass factor > 0.1) have been used as recommended. For further details, including how to apply the averaging kernel and a priori profile in comparisons, data users are encouraged to read the Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) associated with this data product, which are available on https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms.

8.3.2 Status of validation

This section presents a summary of the validation results obtained with the Validation Data Analysis Facility (VDAF) of the S5P Mission Performance Centre (MPC) and by the SSP Validation Team (S5PVT) AO projects CESAR and NIDFORVAL. Up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal at http://mpc-vdaf.tropomi.eu. The status of the FRM data streams is as follows:
• Comparisons of S5P HCHO with UV-Vis MAX-DOAS. At the present stage of S5P routine operation, with the recent addition of the Mohali station (India), six MAX-DOAS stations contribute data routinely to the VDAF Automated Validation Server with a temporal coverage of collocations from May 2018 to February 2022. A problem was identified with the UV channel of the MAX-DOAS at Xianghe, affecting data from November 2021 onwards; these data have been removed from EVDC.

• Data from nine stations are available through the NIDFORVAL AO project. It covers the period from May 2018 to February 2022, dependant on the station.

• Comparisons of S5P HCHO with NDACC FTIR follow the methodology of Vigouroux et al. (2020). The current number of stations is 27 and covers the period May 2018 - November 2021.

8.3.3 Bias

The following results, using ground-based FTIR, MAX-DOAS, and OMI satellite data, show that the TROPOMI HCHO bias is usually well within the 40% mission requirements and always within the 80% upper limit.

8.3.3.1 Fourier Transform Infrared Spectrometers

TROPOMI shows a positive bias of +28±4% for clean stations (HCHO <2.5 Pmolec/cm²) and a negative bias of -29±1% for high emission stations (>8 Pmolec/cm²) in comparison to correlative data from 28 NDACC FTIR stations, covering the period from May 2018 to November 2021, as illustrated in Figure 55. Details about the applied methodology are described in Vigouroux et al. (2020). Using the robust Theil-Sen estimator to derive slope and intercept of TROPOMI vs FTIR, a constant positive bias of 1.06±0.03 Pmolec/cm² and a proportional bias of 0.65±0.02 is obtained (updated from Vigouroux et al. (2020), Fig. 4).

Figure 55: Percent bias between S5P L2_HCHO and NDACC FTIR HCHO column data at each station as a function of the mean FTIR total column (10¹⁶ molec/cm²). The grey bars are the systematic uncertainty on the difference, and the coloured error bars are the 2-σ errors on the bias. The temporal range is May 2018 to November 2021 (updated from Vigouroux et al. (2020), Fig. 3).
8.3.3.2 MAX-DOAS UV-visible Spectrometers

Figure 56 shows difference time series S5P-MAXDOAS for the 6 stations from the AVS, with indication of the 30-day rolling monthly mean or median. Biases, if they occur, are mostly negative, and within the 80% requirements. Pixel size switch and version change (v1 to v2.1, July 2020; v2.1 to v2.2, July 2021, to v2.3, March 2022) have no impact on the comparison.

Figure 56: Time series of the difference [Pmolec/cm²] and relative difference [%] between S5P RPRO+OFFL and MAX-DOAS HCHO column data at six stations. Data was obtained from the VDAF Automated Validation Server on 2022/05/24. Dashed line: pixel size switch at 2019/08/06. Full vertical line: OFFL processor version change v1 → v2.1 on 2020/07/13. Dotted line: v2.2 on 2021/07/01, Green line: v2.3 on 2022/03/06.

The VDAF comparisons have been done so far for a single pixel of S5P versus 0.5h means of MAX-DOAS. This leads to larger scatter. An area averaged 20km column of S5P data in comparison to 3h MAX-DOAS means are better suited to detect seasonal cycles. This is shown in Figure 56 for the station Mohali. Bias, dispersion, and negative values are reduced for the area averaged comparisons.
Figure 57: Time series of HCHO columns [Pmolec/cm$^2$] from S5P RPRO+OFFL and MAX-DOAS data at the Mohali station. (Left) Single S5P pixel, ±0.5h MAX-DOAS averaging, (Right) 20 km area average for ≥10 pixels, ±3h averaging. Data processed 2022/02/28.

In the NIDFORVAL AO project, eight stations are used that provide data in the GEOMS format as required in the project. Figure 58 (upper panel) shows that the median bias varies between -5% to -59%, with a median for all stations of -37%. Among these 8 stations, 3 are providing profiles and averaging kernels (AK) which allow to take the difference in a-priori profiles and vertical resolution of the instruments into account (Rodgers and Connor, 2003), as done with the FTIR network. This is particularly important for the MAX-DOAS instruments due to different shape of AK compared to TROPOMI (see Figure 59). For these three stations, the biases improve when using the averaging kernel (Figure 58, lower panel).

Figure 58: Bias at each station (in %) as a function of the mean DOAS total columns ($10^{16}$ molec/cm$^2$). The grey bars are the systematic uncertainty on the differences, and the coloured error bars are the 2-σ error on the bias. Top panel: DOAS is the normal product without any modification. Bottom panel: DOAS data after Rodgers and Connor (2003) is applied (a priori substitution and smoothing with the TROPOMI averaging kernels).
Further separating the biases for low (< 2.5 Pmolec/cm$^2$) and high (> 8 Pmolec/cm$^2$) HCHO levels as done for FTIR comparisons, we get for the smoothed DOAS comparisons $+27\pm25\%$ and $-10\pm2\%$, respectively. The large uncertainty on the bias for low-HCHO levels is due to the small number of data involved, because the DOAS stations used in this study are not situated in a clean environment (see x-axis in Figure 58, >5 Pmolec/cm$^2$).

8.3.3.3 Consolidation of FTIR and MAX-DOAS validation results on bias

The FTIR data and the MAX-DOAS data (from two streams: NIDFORVAL and VDAF server) used for the S5P HCHO validation are different in scope (FTIR network having more stations and covering a wider range of HCHO values), harmonization (FTIR network being the more harmonized one), vertical sensitivity (FTIR vertical sensitivity being closer to that of TROPOMI) and uncertainty (FTIR having the smaller systematic error and random error uncertainty). Thus, differences between FTIR and MAX-DOAS validation results can at least be partly attributed to the above-cited factors.

The FTIR network covers very low per-station mean HCHO column values (down to 1.2 Pmolec/cm$^2$), which is not the case for the MAX-DOAS network (lower bound is at Uccle, which is moderately polluted). To check the consistency of validation results of both networks one should therefore consider a common range of HCHO levels. De Smedt et al. (2021, paper in preparation), using a larger set of NIDFORVAL MAX-DOAS stations, found that in the HCHO column range of 3-6 Pmolec/cm$^2$, TROPOMI columns do not have a significant bias towards the MAX-DOAS stations, in agreement with the results for FTIR (Vigouroux et al., 2020). Note that the NIDFORVAL MAX-DOAS network does not include mean levels below 3 Pmolec/cm$^2$.

However, one should consider that the results of De Smedt et al. (2021) are taken for unsmoothed MAX-DOAS columns. Also, the previous section makes clear that, unexpectedly, the agreement between unsmoothed MAX-DOAS and smoothed FTIR is in fact better than the agreement between smoothed MAX-DOAS and smoothed FTIR. Note however that (i) there are only 3 smoothed MAX-DOAS in use, and (ii) the result is strongly driven by the MAX-DOAS station Unam with a large bias change upon smoothing. More profile MAX-DOAS data is therefore needed to draw strong conclusions. As the FTIR data have the broadest scope, include more stations and are more harmonized, FTIR validation results are provided as representative quality indicator.
8.3.3.4 OMI QA4ECV comparisons

The TROPOMI HCHO algorithm was designed in parallel with the QA4ECV OMI algorithm in order to create a consistent time series of early afternoon observations. The QA4ECV OMI HCHO dataset is now exceeding 16 years (2005-2020), including three years of overlap with TROPOMI, allowing for a meaningful comparison at different scales. As presented in the TROPOMI HCHO ATBD, all retrieval settings have been chosen as similar as possible for the two L2 products, as well as the auxiliary datasets with the important exception of the cloud products (De Smedt et al., 2018; 2021).

While the QA4ECV OMI product is based on the O2–O2 absorption feature around 477 nm, and considers a fixed cloud albedo of 0.8 (version 2.0), the TROPOMI product uses the S5P operational cloud product in CRB (Cloud as Reflecting Boundary) mode (OCRA/ROCINN-CRB). The S5P ROCINN algorithm is based on the O2 A-band around 760 nm and simultaneously retrieves the cloud-top height and cloud albedo. Systematic differences between the cloud parameters will result in differences in the air mass factors, influencing the comparisons. To get around this difference between OMI and TROPOMI, it is advised to replace the cloud-corrected AMFs by clear-sky AMFs (no cloud correction applied). Both types of AMFs are provided in both L2 products.

We calculated averaged columns in 35 regions covering a broad range of emission levels and observation conditions (large black boxes on Figure 57). As the regions are large, many observations are included (on average 500/day for OMI, 12500/day for TROPOMI). To obtain daily and monthly comparison pairs, we keep coincident days of observations.

Figure 60 presents the mean bias between OMI and TROPOMI HCHO tropospheric columns for the 35 regions. Numbers are provided for daily averaged columns applying a cloud correction (upper panels) or not (lower panels). As discussed in De Smedt et al. (2021), biases up to 30% related to the cloud correction are observed over Tropical regions where the clouds are the highest in altitude (Africa, South America, South Asia), and a smaller but systematic effect, up to 15%, is observed over mid-latitude polluted regions such as China, India, US or Europe. We note that the differences between N_v and N_v_clear are mainly significant for the OMI HCHO columns.
It has been reported that the cloud pressures retrieved from TROPOMI and from OMI present a bias (OMI clouds are higher in altitude, Compernolle et al., 2020). This translates into OMI cloud-corrected air mass factors generally smaller than TROPOMI AMFs by 5 to 30%, depending on the cloud altitude, and therefore in a positive bias of the OMI HCHO VCD compared to the TROPOMI product. It is therefore important to keep in mind that the use of different cloud products may introduce inconsistencies, which may be resolved by using clear HCHO VCDs ($N_{\text{v\_clear}}$).
When comparing $N_{v\text{, clear}}$, the biases are strongly reduced below 10% in all regions where the HCHO levels are larger than 5 Pmolec/cm$^2$, and the TROPOMI columns are found to be slightly larger than OMI on average (-3±1.2%). In mid-Northern-latitudes/moderate emissions (2-5 Pmolec/cm$^2$) regions such as Europe, Central and Western US, North Western Canada, Siberia or Tibet, OMI columns present a remaining bias of about 15±3%, while in the regions of Canada and Alaska, a larger bias of about +30±7% remains. The mean columns are lower over those regions, and differences in sampling (pixel size and OMI row degradation) start playing a bigger role.

### 8.3.4 Dispersion

The dispersion is evaluated for several scenarios: single pixel comparisons with MAX-DOAS (from VDAV-AVS), 20-km radius pixel average comparisons with MAX-DOAS (from NIDFORVAL), and 20-km radius pixel average comparisons with FTIR, for all stations and for clean stations only. The dispersion difference obtained for the 20-km radius pixel averages is also recalculated to a theoretical single-pixel value by multiplying with $\sqrt{(#\text{pixels})}$. However, one should take into account that this formula assumes that random error is uncorrelated and only originates from the satellite.

#### 8.3.4.1 Fourier Transform Infrared Spectrometers

In the current update of the NIDFORVAL project, using now 28 FTIR stations, the median absolute deviation (MAD) remains close to the mission requirement of 12 Pmolec/cm$^2$. In this work, we do not use a single TROPOMI pixel (as in the MPC Automated Validation Server) but an average of about 30 TROPOMI pixels (20 km around the station). The MAD for the 28 stations taken together is 2.49 Pmolec/cm$^2$, corresponding to a theoretical single-pixel dispersion of 2.49×$\sqrt{(#\text{pixels})}$=15 Pmolec/cm$^2$, slightly above the 12 Pmolec/cm$^2$ requirement.

However, to evaluate the TROPOMI precision, it is more relevant to compare the MAD obtained at clean stations only because MAD is less sensitive to the additional collocation error in regions far from emissions. At clean conditions, the TROPOMI precision is much better than the pre-launch requirements: 1.46 Pmolec/cm$^2$ for the 36-pixels-average, corresponding to a theoretical single pixel precision of 9 Pmolec/cm$^2$. These updated results are very similar to results in Vigouroux et al. (2020). We note that, while the pre-launch requirements are reached, and while the provided TROPOMI random uncertainty agree with the estimated single-pixel precision, the MAD of the 20km-averaged-pixels at clean stations is about 1.6 times larger than the random uncertainty budget provided in the TROPOMI files, pointing to a possibly too optimistic TROPOMI random uncertainty budget (Vigouroux et al., 2020).

#### 8.3.4.2 MAX-DOAS UV-visible Spectrometers

The MAD of the difference of S5P (single pixel) with respect to MAX-DOAS ranges from 8 Pmolec/cm$^2$ at Uccle to 10 Pmolec/cm$^2$ at Xianghe. This is within the mission requirement of precision of 12 Pmolec/cm$^2$. Using the 20km-averaged-pixels within NIDFORVAL (~about 42 pixels here), as done for FTIR and not applying vertical smoothing (Rodgers and Connor, 2003), we obtain a MAD of 3.0 and 2.9 Pmolec/cm$^2$ at Cabauw and De Bilt, respectively. This corresponds to a single pixel precision of 19-20 Pmolec/cm$^2$, which is twice larger than the pre-launch requirement of precision. However, if we look at the cleanest DOAS station Uccle to avoid larger collocation errors, the MAD is 2.6 Pmolec/cm$^2$ for the 34-pixels average comparisons, leading to a single pixel precision of 14 Pmolec/cm$^2$. If we apply the Rodgers and Connor (2003) technique, the MAD between TROPOMI and the DOAS$_{\text{smooth}}$ data is reduced at all the five stations, except at Uccle where the smoothing has little effect (Figure 58), leading to a single pixel precision of 15 Pmolec/cm$^2$ there.
In summary, the dispersion of the difference (8 to 10 Pmolec/cm$^2$, mainly polluted stations) is already within the dispersion requirement of 12 Pmolec/cm$^2$ for single-pixel comparisons with MAX-DOAS. Even lower dispersions (2 to 4 Pmolec/cm$^2$) are obtained when using 20-km averaged pixels (NIDFORVAL FTIR, NIDFORVAL MAX-DOAS), as the random error is reduced. However, recalculating to a theoretical single pixel dispersion by multiplying with $\sqrt{(#\text{pixels})}$, the dispersion requirement is now slightly (FTIR) and strongly (MAX-DOAS) higher than the dispersion requirement. But comparison error and FRM random error in the case of MAX-DOAS make an important contribution at polluted stations. The theoretical single pixel dispersion (~8 Pmolec/cm$^2$) at clean FTIR stations is within the dispersion requirement, but 1.6 times larger than the random uncertainty budget provided in the TROPOMI files.

8.3.4.3 Consolidation of FTIR and MAX-DOAS validation results on dispersion

The mission requirement for SSP HCHO on dispersion is 12 Pmolec/cm$^2$. The number is valid at a single-pixel level. The single-pixel comparisons with MAX-DOAS from the VDAF validation server (mainly polluted stations) have a single-pixel dispersion of the difference of 8 to 10 Pmolec/cm$^2$. This is only an upper bound to the SSP HCHO dispersion, as there are also contributions from MAX-DOAS random error and from comparison error. Nonetheless, the result is already within the dispersion requirement of 12 Pmolec/cm$^2$, confirming that at single-pixel level, the dispersion requirement is met.

Even lower dispersions (2 to 4 Pmolec/cm$^2$) are obtained when using 20-km averaged pixels (NIDFORVAL FTIR, NIDFORVAL MAX-DOAS), as the random error is reduced. However, recalculating to a theoretical single pixel dispersion by multiplying with $\sqrt{(#\text{pixels})}$, the dispersion requirement is now slightly (FTIR) and strongly (MAX-DOAS) higher than the dispersion requirement. But comparison error and (in the case of MAX-DOAS) FRM random error make an important contribution at polluted stations. It is therefore more appropriate to focus on clean FTIR stations. The theoretical single pixel dispersion (~8 Pmolec/cm$^2$) at clean FTIR stations is within the dispersion requirement, but 1.6 times larger than the random uncertainty budget provided in the TROPOMI files. We can thus conclude that both at single-pixel level and for 20-km averaged pixel areas, the dispersion requirement is met.

8.3.4.4 OMI QA4ECV Data Record

For individual pixels, the standard deviation of individual OMI and TROPOMI observations in remote regions with no local emissions is about 7 and 5 Pmolec/cm$^2$, respectively. When averaging data over large regions, the dispersion due to random uncertainties is greatly reduced compared to individual observations. As summarized in Table 9, the median absolute deviations of the monthly averaged columns are equivalent for OMI and TROPOMI (1.8 Pmolec/cm$^2$), while the median deviations of their differences are significantly lower (0.5 Pmolec/cm$^2$). This indicates that at this spatiotemporal resolution, the natural variability dominates the dispersion of the averaged observations. Looking at the daily averaged columns, the TROPOMI median deviation is lower than for OMI (2.2/2.7), but still larger than the median deviation of their differences (1.5). Note that these estimates still include the natural variability of the columns themselves. If an in the remote Equatorial Pacific is considered, the observations represent constant background values and the seasonal variability is further reduced. In such conditions, the dispersion of the OMI daily observations is 3.5 Pmolec/cm$^2$, while only 1 Pmolec/cm$^2$ for TROPOMI.

Low dispersion is related to the large number of observations included in the averages. The frequent occurrence of extreme outliers advocates the use of the median difference as a quality indicator instead of the mean difference.
Table 9 - Median absolute deviation of the OMI and TROPOMI daily and monthly averaged columns ($N_{v,\text{clear}}$), in large regions and in 20km-radius area. Median absolute deviations (MAD) of differences between OMI and TROPOMI columns are also given in the last column.

<table>
<thead>
<tr>
<th>Dispersion</th>
<th>OMI MAD [10$^{15}$molec.cm$^{-2}$]</th>
<th>TROPOMI MAD [10$^{15}$molec.cm$^{-2}$]</th>
<th>OMI-TROPOMI MAD [10$^{15}$molec.cm$^{-2}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly Regional</td>
<td>1.8</td>
<td>1.8</td>
<td>0.5</td>
</tr>
<tr>
<td>Daily Regional</td>
<td>2.7</td>
<td>2.2</td>
<td>1.6</td>
</tr>
<tr>
<td>Monthly 20km</td>
<td>3.3</td>
<td>2.5</td>
<td>2.4</td>
</tr>
<tr>
<td>Daily 20km</td>
<td>7.8</td>
<td>3.7</td>
<td>7.1</td>
</tr>
<tr>
<td>Daily 20km in the Equatorial Pacific</td>
<td>3.5</td>
<td>1.0</td>
<td>3.7</td>
</tr>
</tbody>
</table>

8.3.5 Dependence on influence quantities

None to report.

8.3.6 Seasonal and short term variability

Fourier Transform Infrared Spectrometers

The seasonal variability captured by TROPOMI is similar to the one reported by FTIR. As an illustration, Figure 61 shows HCHO time series at Karlsruhe and Hefei. In agreement with the biases discussed in Sect. 7.3.3, TROPOMI shows negative bias when the HCHO columns are the highest (during the July maximum). The correlation between the TROPOMI and FTIR monthly means for all 28 stations is 0.91.

![Figure 61: TROPOMI and FTIR HCHO time series at Karlsruhe and Hefei.](image)

MAX-DOAS UV-Visible Spectrometers

The comparisons of TROPOMI and NIDFORVAL MAX-DOAS HCHO data show a monthly mean correlation of 0.86 for 8 stations, and 0.88 after smoothing for 3 stations providing averaging kernels. But it varies strongly from 0.96 (Xianghe) to 0.55 (UNAM, Mexico City). Time-series at these 2 stations are shown in Figure 62.
OMI QA4ECV comparisons

Day to day correlation between OMI and TROPOMI is very high above emission regions (Figure 62).

Figure 63 shows the comparison of OMI and TROPOMI HCHO columns ($N_{c,\text{clear}}$) averaged over one full year (2019). We observe a very good agreement overall. Differences range from 2 Pmolec/cm² over Tropics to -2 Pmolec/cm² over mid-latitude regions. The gain in TROPOMI precision can be observed at the global scale, mainly at larger latitudes where the OMI sampling is most affected in 2019.
The S5P_L2_HCHO columns are seasonally averaged for spring (March-May 2018) and summer (June-August 2018) and compared to GOME-2B (not shown). The comparison shows similar HCHO spatial patterns. GOME-2 also reports similar HCHO columns as TROPOMI in the same regions.

8.3.8 Impact of UPAS processor upgrades up to version 02.03.00

The processor version changes to V02.01.03 (July 2020), V02.01.04 (December 2020), V02.02.01 (July 2021), and V02.03.00 (March 2022) does not impact the quality of the data as shown in the time evolution of the zonal mean HCHO daily columns (Figure 64). Artefacts near the winter hemispheric poles are not detectable from version V02.01.03 onwards.

![Figure 64: S5P TROPOMI HCHO column data [10^{15} molec/cm^2] as a function of day and latitude. The period is from May 2018 to May 2022. Grid box size in latitude direction is 0.5°. The grey vertical lines mark the processor version changes, the black lines the beginning of each year.](image-url)
Figure 65: Time-series of S5P RPRO+OFFL and FTIR HCHO column at four stations (in molec/cm²), and their relative differences (in %). No significant change is observed after the 1st July 2021 (new version 02.02.01).

The impact of version 02.02.01 on validation results can be seen in Figure 65. The time-series are shown for four FTIR stations with no significant changes in the comparisons after the 1st (13th) July 2021. Note that the period 1st-13th July 2021 has been removed from the comparisons as the S5P HCHO OFFL product is not adequate during these days. This known issue can be solved in a RPRO version of the SSP data.

8.4 Equivalence of L2_HCHO NRTI and OFFL products

We demonstrate the closeness of L2_HCHO NRTI and OFFL products at the MAX-DOAS stations De Bilt and Cabauw. NRTI (V01.01.02 to V02.02.01) and OFFL (RPRO V01.01.05 and OFFL V01.01.05 to V02.02.01) L2_HCHO results, each co-located with MAX-DOAS, were obtained from the VDAF Automated Validation Server. A subset of pixels, common to both NRTI and OFFL, was chosen and differences between NRTI, OFFL, and MAX-DOAS determined. The statistical results are summarized in Table 10.
Table 10 - Statistics on the comparison of the common subset of L2\_HCHO NRTI, L2\_HCHO RPRO+OFFL and co-located MAX-DOAS for the stations Cabauw and De Bilt (*: unit of Pmolec cm$^{-2}$). The Automated Validation Server was consulted on 2022/05/04.

<table>
<thead>
<tr>
<th></th>
<th>Cabauw: 743 common co-locations</th>
<th>De Bilt: 712 common co-locations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NRTI vs OFFL</td>
<td>NRTI vs MXD</td>
</tr>
<tr>
<td>Mean(diff)±sem*</td>
<td>0.01</td>
<td>-3.64±0.31</td>
</tr>
<tr>
<td>Median(diff)*</td>
<td>-0.02</td>
<td>-4.21</td>
</tr>
<tr>
<td>Std(diff)*</td>
<td>1.9</td>
<td>8.3</td>
</tr>
<tr>
<td>1/2 IP68(diff)*</td>
<td>0.8</td>
<td>7.5</td>
</tr>
<tr>
<td>Pearson R</td>
<td>0.97</td>
<td>0.29</td>
</tr>
<tr>
<td>Slope</td>
<td>0.99</td>
<td>0.44</td>
</tr>
</tbody>
</table>

8.4.1 Bias

At the MAX-DOAS stations, the bias (both mean and median difference) of L2\_HCHO NRTI vs. L2\_HCHO OFFL is smaller than that of either L2\_HCHO NRTI or L2\_HCHO OFFL with respect to MAX-DOAS (see Table 10). More importantly, the bias of NRTI vs. OFFL is smaller than the standard error on the mean difference of either NRTI or OFFL with respect to MAX-DOAS. The bias differences between NRTI and OFFL are therefore not statistically significant. Similar conclusions are found using FTIR data (Vigouroux et al., 2020). An example of the spatial distribution of daily differences is shown in Figure 66 (2021/11/120205) for the newest V02.02.01. The relative difference is very low (0.0%) with a Pearson correlation coefficient (CCP) of 0.98. Differences are usually within ±0.5 Pmolec/cm$^2$. 
**8.4.2 Dispersion**

Both standard deviation and ½ 68% interpercentile (1/2IP68) of the NRTI/OFFL differences are much smaller than those between either NRTI/MAX-DOAS or OFFL/MAX-DOAS, indicating a much smaller dispersion between NRTI and OFFL. This is also confirmed by the near-unity NRTI/OFFL Pearson R correlation coefficient and slope. These are much lower for both NRTI and OFFL vs MAX-DOAS.
9 Validation Results: L2_SO2

9.1 L2_SO2 products and requirements

This section reports on the validation of the following geophysical variables of the S5P TROPOMI L2_SO2 product identified in Table 1: the sulphur dioxide total column. Validation results are discussed with respect to the product quality targets outlined in Table 3. The NRTI and OFFL processors producing very similar data products, only validation of the L2_SO2 NRTI product is reported hereafter. Subsection 9.4 demonstrates evidence that NRTI and OFFL data do not differ significantly and that their respective validations yield similar conclusions. Verification checks of the UPAS processor upgrade to version 02.01.03 released on 16 July 2020 and later versions show good consistency of the SO2 data product before and after the processor switch, with rather modest effects. The latest upgrade to version 02.03.00, issued in March 2022, will be assessed in a next update of this report.

9.2 Validation approach

9.2.1 Ground-based networks

**Boundary layer pollution (SO2 total)**

S5P TROPOMI L2_SO2 sulphur dioxide column data are compared to ground-based MAX-DOAS UV-visible observations. However, currently the number of available stations in strongly polluted regions is very rare. Outside strongly polluted regions, the SO2 column is below the detection limit of both the MAX-DOAS and satellite measurements. For the validation of the S5P TROPOMI L2_SO2 sulphur dioxide column data MAX-DOAS measurements at Xianghe (China), Greater Noida (India), and Basra (Iraq) were used so far. SO2 column data are also being retrieved from PGN/Pandora measurements: these initial data are still in a demonstration phase and need uncertainty assessment and validation before getting the status of FRM.

**Volcanic plumes (SO2 enhanced)**

S5P TROPOMI L2_SO2 sulphur dioxide column data are compared to MAX-DOAS UV-visible measurements collected from the Network for Observation of Volcanic and Atmospheric Change (NOVAC) [ER_NOVAC]. Because of the strong SO2 concentration gradients in volcanic plumes, the comparison is not performed using the SO2 columns but rather using the derived SO2 fluxes.

9.2.2 Satellites

S5P TROPOMI L2_SO2 sulphur dioxide column data are compared to similar data from EOS-Aura OMI and Suomi-NPP OMPS.

9.2.3 Field campaigns and modelling support

S5P TROPOMI L2_SO2 sulphur dioxide column data are compared to car MAX-DOAS measurements performed in Lahore.
9.2.4 Test of the expectation of zero SO\textsubscript{2} SCDs (within detection limit) outside volcanic plumes and strongly polluted regions

Outside strongly polluted regions and volcanic plumes, the atmospheric SO\textsubscript{2} concentrations are very low and the corresponding SO\textsubscript{2} columns are below the detection limit of SSP TROPOMI. Thus SSP TROPOMI measurements outside strongly polluted regions and volcanic plumes are used to check the consistency of the S5P TROPOMI L2\_SO\textsubscript{2} sulphur dioxide column data with the assumption of SO\textsubscript{2} slant column densities (SCD) of zero. From this test, also the spread of the S5P TROPOMI L2\_SO\textsubscript{2} sulphur dioxide column data is quantified.

9.3 Validation of L2\_SO\textsubscript{2} NRTI

9.3.1 Recommendations for data usage followed

The quality of the observations depends on many factors which are taken into account in the definition of the qa\_value. While it is a handy way of filtering observations of less quality, the “quality assurance value” should also be considered with caution, as it is a compromise to take into account several aspects, such as: processing errors, presence of clouds or snow/ice, observations affected by sun glint, South Atlantic Anomaly, possible contamination by volcanic SO\textsubscript{2}, absence of background correction, and important variables out of range (importantly the AMF).

The qa\_value is a continuous variable, ranging from 0 (error) to 1 (all is well). In order to avoid misinterpretation of the data quality, it is recommended at the current stage to only use those TROPOMI pixels associated with a qa\_value above 0.5.

For further details, data users are encouraged to read the Product Readme File (PRF), Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) associated with this data product, all available on https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/algorithms

9.3.2 Status of validation

So far the validation of the S5P TROPOMI L2\_SO\textsubscript{2} sulphur dioxide column data is mainly based on satellite to satellite comparisons (Figure 67, Figure 68), for which good agreement is found with OMI and OMPS measurements. Validation for polluted regions using ground based MAX-DOAS data is limited to two stations in polluted regions (Xianghe, China, Greater Noida, close to New Delhi, India, Basra, Iraq, Mexico City, and Wakkerstroom, South Africa, see Figure 69 to Figure 74) and to one field campaign in Lahore (Pakistan). Also here in general good agreement was found. However, it should be noted that for these comparisons the SO\textsubscript{2} columns were mostly close to or below the detection limit of S5P TROPOMI.

S5P TROPOMI L2\_SO\textsubscript{2} sulphur dioxide column data were also compared to ground based MAX-DOAS measurements from the NOVAC network. However, the SO\textsubscript{2} columns were not compared directly, because of the strong gradients across volcanic plumes. Instead the derived SO\textsubscript{2} fluxes were compared, for which good agreement was found.

Outside strongly polluted regions and volcanic plumes, the atmospheric SO\textsubscript{2} SCDs were found to be consistent with the assumption of zero within the measurement uncertainties.
From these comparisons (details are shown below) the following conclusions are drawn:

- over polluted regions the requirements are fulfilled;
- over volcanic plumes the bias requirement is fulfilled, while the random requirement can be exceeded occasionally, which is not seen as a substantial restriction of the data quality;
- from the time series of averaged SO$_2$ SCDs (and their errors and standard deviations) it is concluded that the requirements are fulfilled. The bias and spread are typically below 0.2 DU.

Verification checks of the UPAS processor upgrade to version 02.01.03 released on 16 July 2020 (and subsequent versions 02.01.04 and 02.02.01) show good consistency of the SO$_2$ data product before and after the processor switch, with a visible but rather modest effect (seen in Figure 75 and Figure 77).

![Figure 67](image.png)

**Figure 67:** Top: Comparison of the average distribution (01 Jan 2019 – 31 Dec 2019) of the SO$_2$ VCDs derived from TROPOMI and OMI over regions with strong air pollution. Both data sets show very good agreement. Bottom: Correlation plots TROPOMI versus OMI over the Middle East and India. Note that a fixed AMF of 0.4 was used for both retrievals to exclude the effect of different profile assumptions. Courtesy of Nicolas Theys, BIRA-IASB.
Figure 68: Comparison of TROPOMI and OMPS measurements of the volcanic plume of Kilauea on 26 June 2018. The large figure shows the original TROPOMI data. The two small figures show the spatially degraded TROPOMI data and the OMPS data. The figure right shows the correlation plot of the degraded TROPOMI data versus the collocated OMPS data. Courtesy of C. Li and N. Krotkov, NASA/GSFC.

Figure 69: Comparison of TROPOMI SO$_2$ VCDs to MAX-DOAS measurements at Greater Noida (close to New Delhi, India). The following selection criteria were applied: distance < 15km, CF<0.2, AMF>0.2, MAX-DOAS +/- 1h around S5P overpass. Courtesy of M. Sharma (Sharda University, India) S. Donner, S. Dörner, T. Wagner (MPIC), N. Theys (BIRA-IASB).
Figure 70: Comparison of TROPOMI SO₂ VCDs to MAX-DOAS measurements (daily means) at Xianghe (China). The following selection criteria were applied: distance < 25km, CF<0.2, AMF>0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB).

Figure 71: Comparison of TROPOMI SO₂ VCDs to MAX-DOAS measurements (daily means) at Basra (Iraq) from April 2019 to March 2020. The following selection criteria were applied: distance < 25km, CF<0.2, AMF>0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB), data provided by Nayyef Almaliki, Mustafa Aldossary, Ali Almasoudii, Sebastian Donner, Steffen Dörner, Thomas Wagner.

Figure 72: Comparison of TROPOMI SO₂ VCDs to preliminary Pandora measurements (daily means) at Mexico City from September 2019 to August 2021. The following selection criteria were applied: distance < 25km, CF<0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB), Martin Tiefengraber and Alexander Cede (Luftblick). Pandora data are still in demonstration phase.
9.3.3 Bias

The bias is well within requirements for boundary pollution. From the time series of averaged SO₂ SCDs it is estimated that the bias is within 0.2 DU. For volcanic plumes, very good agreement with other satellite observations is found (<10%), but due to the lack of validation by ground-based measurements, the true bias might be larger in some cases.

9.3.4 Dispersion

The dispersion is well within requirements for observations of the boundary pollution. For observations of dense volcanic plumes, the dispersion is probably within the requirements, but due to the lack of validation by ground-based measurements, the true dispersion is at the moment difficult to quantify. From the time series of the standard deviation of the SO₂ SCDs it is estimated that the dispersion is within 0.2 DU.
Figure 75: Temporal evolution of the measurement error (left) and the standard deviation (right) for selected 5° latitude bands and 3 detector rows from December 2018 to April 2022. Good qualitative agreement between both quantities is found indicating that the random uncertainty is well characterized by the measurement error. Larger errors (and standard deviations) are found at the edges of the detector and towards high latitudes. The jump in the RMS and standard deviations in low and mid-latitudes in August 2019 are caused by the reduction of the ground pixel size. Courtesy of N. Theys (BIRA-IASB).

Figure 76: Temporal evolution of the averaged SO₂ SCD for selected 5° latitude bands and 3 detector rows from December 2018 to April 2022. The values are close to zero and show relatively small day-to-day variations. The larger variations in August 2019 are caused by strong volcanic eruptions. Courtesy Nicolas Theys (BIRA-IASB).
9.3.5 Dependence on influence quantities

Slightly larger bias and dispersion are found towards higher SZA.

9.3.6 Short term variability

The short term variability can be estimated from the time series of averaged SO$_2$ SCDs (outside periods with strong volcanic eruptions). It is estimated to below about 0.1 DU.

9.3.7 Geographical patterns

Slightly larger bias and dispersion are found at higher latitudes, likely as an effect of high solar zenith angles.

9.3.8 Other features

None to report.

9.4 Equivalence of L2_SO2 NRTI and OFFL products

The NRT and offline SO$_2$ products are very similar, as illustrated by the comparison of the SO$_2$ SCDs of both data versions hereafter. Thus, the validation activities performed for the OFFL data product (see above) are also representative for the NRTI data product.

Figure 77: Comparison of the NRT (left) and offline (right) SO$_2$ data products. Shown are the time series of background corrected SO$_2$ SCDs for all 450 detector rows from June 2018 to April 2022. Most of the short time features (vertical colored lines) are caused by individual strong volcanic eruptions. For the NRT product, a strong spike occurs in early April 2020, which is caused by a downlink issue that corrupted an irradiance file. The effect of the UPAS processor switch to version 02.01.03 appears but remains modest compared to other features. Later updates have almost no effect. Courtesy of Nicolas Theys, BIRA-IASB.
10 Validation Results: L2_CO

10.1 L2_CO products and requirements

This section reports on the validation of the following geophysical variables of the S5P TROPOMI L2_CO product identified in Table 1: the carbon monoxide total column. Validation results are discussed with respect to the product quality targets outlined in Table 3. The NRTI and OFFL processors used different approaches up to the NRTI processor version 01.03.02 (implemented on July 3 2019) and the equivalence of the two S5P products is demonstrated in Section 10.4. Since processor version 02.02.00 implemented on July 1 2021, a destriped CO column is added in the OFFL L2 data. Comparison results of L2 CO OFFL standard and destriped products are shown in Section 10.6.

10.2 Validation approach

10.2.1 Ground-based networks

S5P TROPOMI L2_CO carbon monoxide column data are routinely compared to reference measurements obtained from FTIR spectrometers performing network operation in the context of the Network for the Detection of Atmospheric Composition Change (NDACC, http://ndacc.org), the Total Carbon Column Observing Network (TCCON, https://tccondata.org) and the Collaborative Carbon Column Observing Network (COCCON, https://www.imk-asf.kit.edu/english/COCCON.php). Figure 78 displays the geographical distribution of the NDACC and TCCON stations. Near-infrared TCCON measurements provide CO column averaged (xCO) data with typical uncertainty values of 2% for the bias and 1% for the precision. The COCCON measurements are calibrated to TCCON and show similar performance as TCCON. Solar infrared NDACC measurements provide CO total column data with a typical total uncertainty of 3%.

Figure 78: Geographical distribution of NDACC and TCCON FTIR stations measuring atmospheric carbon monoxide column data. Some stations contribute to the two networks.

10.2.2 Satellites

None for this report.
10.2.3 Field campaigns and modelling support

None for this report.

10.3 Validation of L2_CO OFFL

10.3.1 Recommendations for data usage followed

The Product Readme File (PRF) recommends the use of only S5P data with a qa_value above 0.5. The validation results reported hereafter are obtained by filtering the pixels using the parameters mentioned in the PRF, distinguishing three cases based on cloud filtering:

1. Clear sky: cloud height below 500 m and cloud optical depth below 0.5 (qa_value=1);
2. Cloud: cloud height below 5000 m and cloud optical depth above 0.5 (qa_value=0.7);
3. All: cloud height below 5000 m.

For further details, data users are encouraged to read the Product Readme File (PRF), Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) associated with this product: https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms.

10.3.2 Status of validation

This section presents a summary of the key validation results obtained by the MPC VDAF and by SSPVT AO projects. It is based on the validation methodology reported at the third SSP Validation Team Workshop (ESA/ESRIN, November 11-14, 2019). Individual contributions to the workshop are archived in https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5-precursor-workshop-2019/sentinel-5p, while up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal at http://mpc-vdaf.tropomi.eu.

Current conclusions are based on the amount of reference measurements available at the time of this analysis, yielding comparison pairs from November 2017 through April 2022. Routine validation is done using the Automated Validation Server of the MPC VDAF, the CO validation system operated at BIRA-IASB, and the HARP toolset.

TROPOMI observations co-located with the TCCON, COCCON and NDACC measurements are found by selecting all filtered TROPOMI pixels within a radius of 50 km around each station and with a maximal time difference of 1h for TCCON and COCCON and 3h for NDACC observations. The 1-hour interval can be justified by noting that TCCON and COCCON instruments acquire only one type of spectra, while NDACC instruments are required to measure different types of spectra, making the number of CO observations sparser. In the comparison procedure with TCCON and COCCON data, the a priori in the TCCON and COCCON retrievals have been substituted with the S5P CO a priori (Rodgers 2003). The validation procedure for the NDACC, COCCON and TCCON based comparisons includes an adaptation of the TROPOMI CO column to the altitude of the ground-based FTIR instrument.

Since August 6, 2019 (orbit 9388), S5P measures with increased spatial resolution from 7km to 5.5km along track. This change in operations did not change the performance of the CO NRTI and OFFL product. TCCON released a new data version (GGG2020) on April 26, 2022. No recent data has been submitted for TCCON with the old processor. Therefore, the TCCON sections have not been updated in this report and will be updated in the next report with the latest GGG2020 data.
10.3.3 Bias

The systematic difference between S5P L2_CO daily mean data and correlative ground-based measurements is on an average 6.5% with respect to NDACC data, 9.84% with respect to TCCON data, and 4.8% with respect to COCCON data. At some stations like mountain stations, those values are exceeded likely because of geographical colocation issues. This bias estimate falls well within the mission requirements. Figure 80 - Figure 82 show the biases over the period Nov 2017– April 2022 sorted by latitude. No significant latitudinal dependence of the bias is observed. Figure 79 does not show any significant degradation in bias with time (note that the longer period covers different processor versions). Figure 79 also shows a slight increase of bias during local winter.

The processor update in July 2021 (version 02.02) introduces a different spectroscopy, which leads to a change in the overall bias. First results using the rapid delivery data of NDACC indicate that the systematic difference reduces from 6.5% (before July 2021) to just below 3% (after July 2021) (see also §10.5).
Figure 79: Relative bias between S5P L2_CO OFFL and ground-based CO column data at NDACC (top), TCCON (middle), and COCCON (bottom) FTIR stations. Over the Nov. 2017 – April 2022 time period the plots do not show a clear meridian dependence or temporal change in the weekly averaged biases. No recent measurements are available for the TCCON comparison. The vertical black line indicates the implementation date of the processor update 02.02, which introduces a change in spectroscopic parameters.
Figure 80: Bar chart of the relative mean difference between SSP L2, CO OFFL & NRTI and FTIR CO column data at 23 NDACC stations, for all data within the time range from November 2017 until April 2022 showing RPRO/OFFL. The stations are sorted with decreasing latitude. All biases are below 15% except at Altzomoni (a mountain city near Mexico City). Both NRTI (since July 2019) and OFFL behave similar w.r.t. systematic biases.
Figure 81: Bar chart of relative mean difference between SSP L2_CO OFFL and FTIR CO column data at 29 TCCON stations for all data within the time range Nov 2017 until May 2021. The stations are sorted with decreasing latitude. The majority of the OFFL biases are below 12% except in the Arctic and mountain station (Izaña, Nicosia, and Garmisch) where the bias is slightly above 12%. Xianghe station lies in a polluted region where we see almost zero bias. No recent measurements are available for the TCCON comparison.
10.3.4 Dispersion

The 1σ dispersion of the relative mean bias around its mean is of the order of 5%. The individual values for the different stations are indicated in Figure 80 - Figure 82. This dispersion can be considered as an upper boundary of the random uncertainty of the satellite data.

10.3.5 Dependence on influence quantities

At this stage, the evaluation of potential dependence of the S5P bias and spread on the Solar Zenith Angle (SZA) shows an increase of the relative bias with the solar zenith angle of about 10% between 10deg and 80deg.

Figure 82: Bar chart of relative mean difference between S5P L2_CO OFFL and FTIR CO column data at three COCCON stations for all data within the time range Nov 2017 until April 2022. The stations are sorted with decreasing latitude.
Figure 83: Relative difference (daily mean) between SSP L2_CO RPRO/OFFL and NDACC (top), TCCON Sodankylä (middle), and COCCON (bottom) carbon monoxide total column as a function of the TROPOMI solar zenith angle, in the ‘all’ case. The Wollongong time series contains some outlying values, which can be attributed to co-location mismatches during the Australian fires in Jan 2020 and are not shown in the NDACC plot.
10.3.6 Short term variability

For all the NDACC, TCCON, and COCCON stations, short scale temporal variations in the CO column as captured by ground-based instruments are reproduced very similarly by S5P L2_CO OFFL. This overall good agreement is confirmed by individual Pearson correlation coefficients well above 0.6 for all sites and on average reaching 0.9 (Figure 84).
Figure 84: Taylor diagrams for daily mean differences between SSP L2_CO OFFL and ground-based networks CO data: NDACC (top), TCCON (middle), and COCCON (bottom) for our all case of pixel selection criteria (Boulder and Ascension cover a limited period), NDACC Wollongong has reduced correlation due to co-location mismatches during the Australian fires in Jan 2020).
10.3.7 Geographical patterns

Individual S5P L2_CO CO column data show stripes of erroneous CO values < 5% in the flight direction, probably associated with calibration issues of TROPOMI, see Figure 85 (top) below. This data quality issue is known but not covered by the quality flags, and should be kept in mind when looking at the carbon monoxide data product and at preliminary validation results. Since the implementation of processor 020202 in July 2021, a de-striping algorithm is implemented in the OFFL product. Both standard and de-striped CO corrected column values are available for each pixel.

Figure 85: Example showing CO column data on July 2 2021 above Europe, comparing L2_CO NRTI CO column data (top) using processor 010400 and L2_CO OFFL CO corrected (destriped) columns (bottom) from processor 02.02.00.
TROPOMI CO column data also suffer from instrumental effects of the South Atlantic Anomaly (SAA), see Figure 86.

Figure 86: S5p OFFL xCO pixels measured on August 1, 2019, over South America and the Atlantic Ocean. Outlying pixels occurs (including negative values) in the South Atlantic Anomaly.

10.3.8 Other features

NRTI granules from one S5P orbit have overlapping pixels. In order to avoid duplicated pixels in the validation statistics, pixels from the first 12 (before July 3 2019) or 16 (after July 3 2019) scanlines have been filtered.
10.4 Equivalence of L2_CO OFFL and NRTI products

On July 3, 2019, the L2_CO NRTI processor changed to use the same settings as the OFFL processor. Figure 87 confirms that the statistical quality indicators for both OFFL and NRTI since the processor change are very similar.

Figure 87: Comparison of relative biases against TCCON CO column data for the S5P L2_CO OFFL and NRTI data versions, from July 3 2019 through May 2021. The quality of both data sets is similar. Over this period the OFFL processor has produced data with a relative mean bias of 10.05% ± 5.74% and a correlation coefficient 0.93 with respect to TCCON data; the NRTI processor has produced data with similar statistics: a relative mean bias 9.92% ± 5.74% and a correlation 0.93.

10.5 Accuracy estimate for processor 02.02

On July 1, 2021 the OFFL processor was updated to use different spectroscopic parameters (see the S5P CO Product Readme File). This section provides a first estimate of the accuracy for processor 02.02. From Table 11, the impact of the update in spectroscopic parameters is observed in change in the relative bias: from 6.3% for processor 01 prior to July 2021 to 2.9% for processor 02. The latter estimate should be considered preliminary and must be confirmed with consolidated NDACC data, when a longer time series is available and with TCCON data.
Table 11 NDACC network statistics comparing the OFFL data with processor version 02 (since July 1 2021) against the processor version 01 (Nov 2017-June 2021). The update in July introduced new spectroscopic parameters, which have a systematic effect and this is seen in the bias estimate. The estimate uses the rapid delivery NDACC data.

<table>
<thead>
<tr>
<th>station</th>
<th>S5P-NDACCOFFL v02</th>
<th>S5P-NDACCOFFL v01</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rel. std</td>
<td>correl.</td>
</tr>
<tr>
<td>THULE</td>
<td>315</td>
<td>0.99</td>
</tr>
<tr>
<td>KIRUNA</td>
<td>77</td>
<td>0.99</td>
</tr>
<tr>
<td>KARLSRUHE</td>
<td>22</td>
<td>1.1</td>
</tr>
<tr>
<td>JUNGFRAUJOCH</td>
<td>242</td>
<td>1</td>
</tr>
<tr>
<td>TORONTO</td>
<td>282</td>
<td>1.1</td>
</tr>
<tr>
<td>BOULDER.CO</td>
<td>1100</td>
<td>1.2</td>
</tr>
<tr>
<td>IZANA</td>
<td>79</td>
<td>0.92</td>
</tr>
<tr>
<td>MAUNA.LOA.HI</td>
<td>179</td>
<td>1.1</td>
</tr>
<tr>
<td>LA.REUNION.MAIDO</td>
<td>136</td>
<td>1.2</td>
</tr>
<tr>
<td>WOLLONGONG</td>
<td>355</td>
<td>1.1</td>
</tr>
<tr>
<td>LAUDER</td>
<td>676</td>
<td>0.9</td>
</tr>
<tr>
<td>ARRIVAL.HEIGHTS</td>
<td>126</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>1.05</td>
<td>0.93</td>
</tr>
</tbody>
</table>

10.6 Comparison of L2_CO OFFL standard and destriped products

The processor update to version 02.02.00 on July 1 2021 introduced a de-striped CO column in the S5P OFFL L2 data. A comparison (Table 12) between the standard and de-striped column using the “closest” pixel comparison shows that the de-striped product has reduced dispersion in the relative differences (see also the results on the VDAF server). Relative bias and correlation do not show significant changes.

Table 12 NDACC network statistics comparing the OFFL destriped CO columns against the standard CO columns available in the S5P L2 files since July 2021. The dispersion on relative differences is reduced with approx. 1.5% in the destriped product.

<table>
<thead>
<tr>
<th>station</th>
<th>S5P-NDACCOFFL destriped</th>
<th>S5P-NDACCOFFL standard</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>rel. std</td>
</tr>
<tr>
<td>THULE</td>
<td>320</td>
<td>1</td>
</tr>
<tr>
<td>KIRUNA</td>
<td>70</td>
<td>0.94</td>
</tr>
<tr>
<td>BREMEM</td>
<td>5</td>
<td>1.2</td>
</tr>
<tr>
<td>KARLSRUHE</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>GARMISCH</td>
<td>532</td>
<td>0.9</td>
</tr>
<tr>
<td>ZUGSPITZE</td>
<td>250</td>
<td>1.1</td>
</tr>
<tr>
<td>JUNGFRAUJOCH</td>
<td>242</td>
<td>0.8</td>
</tr>
<tr>
<td>TORONTO</td>
<td>289</td>
<td>1.1</td>
</tr>
<tr>
<td>BOULDER.CO</td>
<td>1106</td>
<td>1</td>
</tr>
<tr>
<td>IZANA</td>
<td>84</td>
<td>0.9</td>
</tr>
<tr>
<td>MAUNA.LOA.HI</td>
<td>179</td>
<td>0.6</td>
</tr>
<tr>
<td>LA.REUNION.MAIDO</td>
<td>137</td>
<td>0.7</td>
</tr>
<tr>
<td>WOLLONGONG</td>
<td>357</td>
<td>0.9</td>
</tr>
<tr>
<td>LAUDER</td>
<td>677</td>
<td>0.8</td>
</tr>
<tr>
<td>ARRIVAL.HEIGHTS</td>
<td>127</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>0.94</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>0.91</td>
<td>0.79</td>
</tr>
</tbody>
</table>
11 Validation Results: L2_CH4

11.1 L2_CH4 products and requirements

This section reports on the validation of the following geophysical variables of the S5P TROPOMI L2_CH4 product identified in Table 1: the methane total column. Validation results are discussed with respect to the product quality targets outlined in Table 3.

11.2 Validation approach

11.2.1 Ground-based networks

S5P TROPOMI L2_CH4 methane column data are routinely compared to reference measurements obtained from FTIR spectrometers performing network operation in the context of the Network for the Detection of Atmospheric Composition Change (NDACC, http://ndacc.org), the Total Carbon Column Observing Network (TCCON, https://tccondata.org) and the Collaborative Carbon Column Observing Network (COCCON, https://www.imk-asf.kit.edu/english/COCCON.php). Figure 78 displays the geographical distribution of the NDACC and TCCON stations. Near-infrared TCCON measurements provide calibrated methane column averaged (xCH4) data with typical uncertainty values of 0.5% for the precision and 0.2% for the accuracy. The COCCON measurements are calibrated to TCCON and show similar performance as TCCON. Solar infrared NDACC measurements provide CH4 total column data with a lower accuracy (typically 3%) and precision (1.5%). The required accuracy (<1.5%) and precision (<1%) for S5P implies that we mainly focus on the validation with TCCON and COCCON measurements.

11.2.2 Satellites

XCH4 measurements by the Thermal and Near Infrared Sensor for Carbon Observation Fourier transform spectrometer (TANSO-FTS) on board the Greenhouse gases Observing SATellite (GOSAT) satellite are used for the validation of the TROPOMI XCH4 data. GOSAT was launched in 2009, and it performs three-point observations in a cross-track swath of 790 km with 10.5 km resolution on the ground at nadir, which results in global coverage approximately every 3 days.

We use the GOSAT proxy XCH4 data product produced at SRON in the context of the ESA GreenHouse Gas Climate Change Initiative (GHG CCI) project (Buchwitz et al., 2019, 2017). This XCH4 product is retrieved using the RemoTeC/proxy retrieval algorithm. The proxy approach (Frankenberg et al., 2005) infers a CO2 and CH4 total column from observations at 1.6 µm, ignoring any atmospheric scattering in the retrieval.

To compare TROPOMI and GOSAT XCH4, we compute daily mean XCH4 in a 2 x 2 degree grid, and then estimate the average bias and its standard deviation.
Figure 88 shows XCH4 measured by GOSAT, TROPOMI, and the ratio of both. Over the ocean, the comparison yields a mean bias of -4.4 ppb ± 15.7 ppb (−0.24±0.85%) and a Pearson's correlation coefficient of 0.80. Over land, the comparison results in a bias before correction of -13.8 ± 16.1 ppb (−0.7 ± 0.8 %) and a Pearson's correlation coefficient of 0.87.

11.2.3 Other TROPOMI XCH4 products

Besides the operational TROPOMI XCH4 product, there is a scientific product using the weighting function modified differential optical absorption spectroscopy (WFM-DOAS) method to retrieve CO and CH4 (Schneising et al., 2019). As part of the activities in the Methane+ project, the pre-operational TROPOMI XCH4 product is compared to the WFMD product.
To compare the two TROPOMI XCH\(_4\) products, we compare daily collocations, and then estimate the average bias and its standard deviation. **Figure 89** shows for 2020 the average XCH\(_4\) retrieved by WFMD-DOAS algorithm and by the pre-operational algorithm, and their ratio. The WFMD-DOAS product covers areas over the ocean further away than the sun-glint limit, and it also does not perform a strict cloud filtering as the operational retrieval. Over the ocean, the comparison yields a mean bias of -2.4 ppb ± 25.9 ppb (−0.14± 1.4 %) and a Pearson's correlation coefficient of 0.56.

11.2.4 Field campaigns and modelling support

None for this report.

11.3 Validation of L2_CH4 OFFL

11.3.1 Recommendations for data usage followed

The Product Readme File (PRF) recommends the use of only S5P data with a qa_value above 0.5.

The S5P L2 data contains two xCH\(_4\) column values: the standard retrieved product and a bias corrected product. Both products are validated separately, but only the bias corrected is mentioned in the quality indicators in **Table 2**.

For further details, data users are encouraged to read the Product Readme File (PRF), Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) associated with this data product: [https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms](https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms).
11.3.2 Status of validation

This section presents a summary of the key validation results obtained by the MPC VDAF and by S5PVT AO projects. The results reported here are an update of those presented and discussed at the 3rd S5P Validation Team Workshop (ESA/ESRIN, November 11-14, 2019). Individual contributions to the workshop are archived in https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5-precursor-workshop-2019/sentinel-5p, while up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal at http://mpc-vdaf.tropomi.eu.

TROPOMI observations co-located with the ground-based FTIR measurements are found by selecting all filtered TROPOMI pixels within a radius of 100 km around each station and with a maximal time difference of 1h for TCCON and COCCON and 3h for NDACC observations. The 1-hour interval can be justified by noting that TCCON and COCCON instruments acquire only one type of spectra, while NDACC instruments are required to measure different types of spectra, making the number of CH₄ observations sparser. In the comparison, the a priori in the TCCON, COCCON and NDACC retrievals have been substituted with the S5P CH₄ a priori (Rodgers 2003). For NDACC data the method described in Rodgers (2003) is followed one step further and the FTIR CH₄ concentration profile (with the S5P prior substituted) is additionally smoothed with the S5P column averaging kernel. The validation procedure for the NDACC, COCCON and TCCON based comparisons includes an adaptation of the TROPOMI CH₄ column to the altitude of the ground-based FTIR instrument.

Current conclusions are based on the S5P and reference measurements available at the time of this analysis, which yield comparison pairs from November 2017 through April 2022. Routine validation is done using the Automated Validation Server of the MPC VDAF, the CH₄ validation system operated at BIRA-IASB, and the HARP toolset, and shows an up-to-date comparison.

Since August 6 2019 (orbit 9388) S5P measures with increased spatial resolution from 7km to 5.5km along track. This change in operations did not change the performance of the methane OFFL product. Since the processor version 02.03.01 update on November 14 2021 retrievals from sun-glint pixels are available, see Section 11.5. TCCON released a new data version (GGG2020) on April 26, 2022. No recent data has been submitted for TCCON with the old processor. Therefore, the TCCON sections have not been updated in this report and will be updated in the next report with the latest GGG2020 data.
11.3.3 Bias

The systematic difference (the mean of all relative differences) between S5P L2_CH4 and TCCON dry air methane column data is on average -0.47% (standard) and -0.03% (bias corrected), well within the mission requirements. Only at a few TCCON stations, the bias is slightly higher than 1.5% for the standard S5P methane column. Comparison against COCCON show an average systematic difference of -0.66 % for the standard methane product and -0.16 % for the bias corrected methane product. Comparisons against NDACC in Table 13 estimate the relative difference of at -0.5% for the standard S5P methane product and +0.3% for the bias corrected product. The NDACC and TCCON reference data are obtained from different spectral regions and explains the difference in bias estimates.

The processor update in July 2021 introduces a change in spectroscopic parameters. Due to the lack of recent TCCON measurements and the sparse presence of rapid delivery NDACC data, the preliminary assessment of the accuracy after July 2021 in section 11.4 should be confirmed with consolidated NDACC and TCCON data.

Figure 90: Mosaic plots of relative biases between S5P L2_CH4 RPRO+OFFL and ground-based CH4 TCCON column data for the bias corrected (top) and standard (bottom) methane products. Over the November 2017 – September 2021 time period the plots do not show a clear meridian dependence or temporal change in the weekly averaged biases. No recent measurements are available for the TCCON comparison.
Figure 91: Mosaic plots of relative biases between S5P L2_CH4 RPRO+OFFL and ground-based CH4 NDACC column data for the bias corrected methane products. Over the November 2017 – May 2022 time period the plots do not show a clear meridian dependence or temporal change in the weekly averaged biases. Processor changes that may affect the performance are indicated by vertical black lines. The introduction of sun-glint pixels introduces results for the Maïdo site on Reunion Island.

Figure 92: Mosaic plots of relative biases between S5P L2_CH4 RPRO+OFFL and ground-based CH4 COCCON column data for the bias corrected (top) and standard (bottom) methane products for the period November 2017 – May 2022.
Figure 93: Chart of relative mean difference between S5P L2_CH4 and FTIR CH4 column data at 26 TCCON stations within the time range November 2017 till May 2021. The stations are sorted with decreasing latitude. The relative mean difference of the standard xCH4 product slightly exceeds the mission requirements (bias below 1.5%) only at a few TCCON stations (i.e. Sodankylä, East Trout Lake, Park Falls and Wollongong). However, it never exceeds the mission requirements for the bias corrected product.
<table>
<thead>
<tr>
<th>Station</th>
<th>#</th>
<th>rel. NDACC std</th>
<th>correlation</th>
<th>rel diff bias (%)</th>
<th>std rel diff (%)</th>
<th>#</th>
<th>rel. NDACC std</th>
<th>correlation</th>
<th>rel diff bias (%)</th>
<th>std rel diff (%)</th>
<th>Lat</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUREKA</td>
<td>309</td>
<td>1.7</td>
<td>0.43</td>
<td>-0.66</td>
<td>1.11</td>
<td>318</td>
<td>1.4</td>
<td>0.56</td>
<td>0.17</td>
<td>1.06</td>
<td>80</td>
</tr>
<tr>
<td>NY.ALESUND</td>
<td>37</td>
<td>1.2</td>
<td>0.65</td>
<td>1.31</td>
<td>0.91</td>
<td>37</td>
<td>1.1</td>
<td>0.74</td>
<td>2.3</td>
<td>0.81</td>
<td>78.9</td>
</tr>
<tr>
<td>THULE</td>
<td>537</td>
<td>1.0</td>
<td>0.66</td>
<td>2.18</td>
<td>1.1</td>
<td>538</td>
<td>0.9</td>
<td>0.73</td>
<td>3.38</td>
<td>1.02</td>
<td>76.5</td>
</tr>
<tr>
<td>KIRUNA</td>
<td>552</td>
<td>0.8</td>
<td>0.53</td>
<td>-2.21</td>
<td>1.15</td>
<td>552</td>
<td>0.8</td>
<td>0.57</td>
<td>-1.18</td>
<td>1.16</td>
<td>67.8</td>
</tr>
<tr>
<td>SODANKYLA</td>
<td>2220</td>
<td>1.2</td>
<td>0.59</td>
<td>-1.96</td>
<td>1.16</td>
<td>2226</td>
<td>1.2</td>
<td>0.64</td>
<td>-0.89</td>
<td>1.11</td>
<td>67.4</td>
</tr>
<tr>
<td>HARESTUA</td>
<td>205</td>
<td>1.1</td>
<td>0.43</td>
<td>-0.39</td>
<td>1</td>
<td>205</td>
<td>1.2</td>
<td>0.49</td>
<td>0.67</td>
<td>0.93</td>
<td>60.2</td>
</tr>
<tr>
<td>ST.PETERSBURG</td>
<td>893</td>
<td>1.1</td>
<td>0.52</td>
<td>-0.97</td>
<td>1.05</td>
<td>893</td>
<td>1.1</td>
<td>0.62</td>
<td>-0.13</td>
<td>0.94</td>
<td>59.9</td>
</tr>
<tr>
<td>BREMEN</td>
<td>206</td>
<td>1.1</td>
<td>0.69</td>
<td>0.3</td>
<td>0.89</td>
<td>206</td>
<td>1.1</td>
<td>0.73</td>
<td>0.83</td>
<td>0.84</td>
<td>53.1</td>
</tr>
<tr>
<td>KARLSRUHE</td>
<td>763</td>
<td>0.8</td>
<td>0.64</td>
<td>-1.63</td>
<td>0.74</td>
<td>759</td>
<td>0.8</td>
<td>0.65</td>
<td>-0.98</td>
<td>0.67</td>
<td>49.1</td>
</tr>
<tr>
<td>GARMISCH</td>
<td>1213</td>
<td>1.1</td>
<td>0.7</td>
<td>-0.63</td>
<td>0.78</td>
<td>1213</td>
<td>1</td>
<td>0.79</td>
<td>0.28</td>
<td>0.68</td>
<td>47.5</td>
</tr>
<tr>
<td>ZUGSPEITZ</td>
<td>863</td>
<td>1.3</td>
<td>0.56</td>
<td>-0.59</td>
<td>1.01</td>
<td>863</td>
<td>1.2</td>
<td>0.65</td>
<td>0.28</td>
<td>0.94</td>
<td>47.4</td>
</tr>
<tr>
<td>JUNGFRAUJOCH</td>
<td>426</td>
<td>1.2</td>
<td>0.48</td>
<td>-0.97</td>
<td>1.2</td>
<td>426</td>
<td>1.1</td>
<td>0.48</td>
<td>-0.25</td>
<td>1.25</td>
<td>46.5</td>
</tr>
<tr>
<td>TORONTO</td>
<td>622</td>
<td>1.8</td>
<td>0.06</td>
<td>0.19</td>
<td>2.49</td>
<td>622</td>
<td>1.7</td>
<td>0.02</td>
<td>0.84</td>
<td>2.59</td>
<td>43.6</td>
</tr>
<tr>
<td>RIKUBETSU</td>
<td>21</td>
<td>1.3</td>
<td>0.16</td>
<td>0.07</td>
<td>1.94</td>
<td>21</td>
<td>1.2</td>
<td>0.09</td>
<td>0.85</td>
<td>2.09</td>
<td>43.5</td>
</tr>
<tr>
<td>BOULDER.CO</td>
<td>2116</td>
<td>1.6</td>
<td>0.73</td>
<td>-0.48</td>
<td>0.88</td>
<td>2116</td>
<td>1.3</td>
<td>0.75</td>
<td>-0.44</td>
<td>0.85</td>
<td>40</td>
</tr>
<tr>
<td>MAUNA,LOA,H1</td>
<td>35</td>
<td>1.2</td>
<td>0.29</td>
<td>-1.07</td>
<td>1.81</td>
<td>35</td>
<td>1.1</td>
<td>0.24</td>
<td>-0.01</td>
<td>1.97</td>
<td>19.5</td>
</tr>
<tr>
<td>ALTZOMONI</td>
<td>395</td>
<td>0.9</td>
<td>0.44</td>
<td>2.98</td>
<td>0.81</td>
<td>395</td>
<td>0.9</td>
<td>0.47</td>
<td>2.85</td>
<td>0.79</td>
<td>19.1</td>
</tr>
<tr>
<td>PORTO.VELHO</td>
<td>68</td>
<td>1.0</td>
<td>0.2</td>
<td>-2.72</td>
<td>0.89</td>
<td>68</td>
<td>1</td>
<td>0.21</td>
<td>-1.51</td>
<td>0.87</td>
<td>-8.8</td>
</tr>
<tr>
<td>LA.REUNION.NAIDO</td>
<td>8</td>
<td>2.7</td>
<td>0.57</td>
<td>-0.51</td>
<td>0.67</td>
<td>8</td>
<td>3</td>
<td>0.73</td>
<td>-0.09</td>
<td>0.63</td>
<td>-21.1</td>
</tr>
<tr>
<td>WOLLONGONG</td>
<td>1413</td>
<td>0.9</td>
<td>0.63</td>
<td>-1.58</td>
<td>0.98</td>
<td>1416</td>
<td>0.9</td>
<td>0.66</td>
<td>-0.73</td>
<td>0.95</td>
<td>-34.4</td>
</tr>
<tr>
<td>LAUERD</td>
<td>1664</td>
<td>1.3</td>
<td>0.69</td>
<td>-0.86</td>
<td>0.95</td>
<td>1667</td>
<td>1.3</td>
<td>0.66</td>
<td>-0.47</td>
<td>1</td>
<td>-45</td>
</tr>
<tr>
<td>ARRIVAL.HEIGHTS</td>
<td>172</td>
<td>1.1</td>
<td>0.45</td>
<td>0.06</td>
<td>1.08</td>
<td>172</td>
<td>0.9</td>
<td>0.71</td>
<td>1.43</td>
<td>0.85</td>
<td>-77.8</td>
</tr>
</tbody>
</table>

Table 13 – Overview of statistical quality indicators for the co-located S5P and NDACC time series. Izaña has only four colocation and is left out from the averaged statistics. The relative mean difference of the standard xCH4 product slightly exceeds the mission requirements (bias below 1.5%) only at a few NDACC stations (i.e. Ny Alesund, Thule and Altzomoni).
11.3.4 Dispersion

The 1σ spread of the relative difference (between the S5P and the TCCON and COCCON methane column data) around the mean value is below the mission requirements (precision <1%) for both the bias corrected and standard products. The individual values for the different stations are indicated in Figure 93.

Because NDACC measurements are reported with a higher random uncertainty, the NDACC estimate for the precision of 1.5% therefore exceeds the mission requirement and should be ignored.

11.3.5 Dependence on influence quantities

At this stage, the evaluation of potential dependence of the S5P bias and spread on the Solar Zenith Angle (SZA) is hard to evaluate: at high latitude stations e.g., Sodankylä and Kiruna, the bias during spring and autumn (both seasons have high SZA) changes sign. At other low latitude stations, we see a SZA dependence of bias e.g. a bias of about 0.5% is seen at Edwards.

The relative differences show a dependence on the surface albedo, which is corrected in the bias corrected product. The relative difference of the bias corrected product shows a remaining weak dependence in low albedo case (which corresponds to the shape and ‘goodness’ of the polynomial fit used to determine the S5P bias correction factor).

Figure 94: Dependence of the S5P-TCCON relative difference on solar zenith angle (top) and surface albedo (bottom). The left column shows the standard S5P product and the right column the bias corrected S5P product. The bias correction removes the surface albedo dependence of the standard S5P product.
11.3.6 Short term variability

For all the NDACC, COCCON and TCCON stations, short scale temporal variations in the CH₄ column as captured by ground-based instruments are reproduced very similarly by S5P L2_CH4 OFFL. The individual Pearson correlation coefficients are on average 0.70, see Figure 95 and Figure 96. At some stations the correlations are very low (e.g. Sodankylä, Darwin). This is probably due to the qa_value filtering which, at some stations, does not filter all bad pixels, see also Section 11.3.8.

**Figure 95**: Taylor diagrams for differences between S5P L2_CH4 RPRO/OFFL and TCCON methane column data: standard (top) and bias corrected (bottom) S5P methane columns. At almost all stations the variability of the S5P column data is higher compared to the variability in the TCCON data.
Figure 96: Taylor diagrams for differences between S5P L2_CH4 RPRO/OFFL and COCCON methane column data: standard (top) and bias corrected (bottom) S5P methane columns.

11.3.7 Geographical patterns

Single TROPOMI overpasses show stripes of erroneous CH$_4$ values in the flight direction (see Figure 97 left). For orbits before orbit 7644 (April 5, 2019) not all pixels above inland water are filtered with the qa_value flag, see Figure 97 (right, above Caspian Sea).
Figure 97: Maps showing XCH₄ concentrations above the Middle East measured by S5P on May 23 2018. The left panel shows all available ground pixels; the right panel shows only pixels with qa_value > 0.5. The left panel shows the presence of stripes in the flight direction and the right panel shows the presence of filtered pixels above inland water (Caspian Sea).

11.3.8 Change in Cloud Mask

Since March 11 2020 the cloud data used for the CH₄ data quality filtering has changed (cf PDGS release V2.6.0). The ECM Cloud Mask is used instead of VICMO since orbit 12432. The qa_value threshold of 0.5 removes significantly more pixels since March 11. This has resulted in a reduced number of co-located days and pixels that are found for the NDACC and TCCON stations. The XCH₄ data earlier than March 11 2020 uses a cloud fraction (CF) pre-filter of 0.02 and a post-filter of 0.001 (cloud free). The definition of CF was changed by the L2_CH4 product developer at SRON such that it uses both confidently clear and probably clear pixels for the ECM Cloud Mask case. Furthermore, the CF thresholds were relaxed and tested against collocated TCCON data. The validation results with the relaxed CF parameter show increased number of pixels in collocation; the bias and precision do not worsen and remain within the mission requirements (see Figure 98). A CF threshold of 0.001 was proposed by the L2_CH4 team and TCCON validation results supports this change. The cloud fraction definition has been optimized in version 02.02.00 implemented since July 1, 2021. Additional important changes implemented in version 02 are the update of the CH₄, CO and H₂O cross sections with improved SEOM-IAS spectroscopy. The a-posteriori bias correction is done independent of any reference data as done for earlier version 01.xxx.
11.3.9 Other features

Filtering on \texttt{qa\_value} > 0.5 does not remove all pixels considered bad. Some pixels with too low and too high methane concentrations are still present (see Figure 99).

Outlying methane values are observed along coastline regions or mountain regions, see for example Greenland in Figure 100.
Figure 99: S5P L2 CH4 XCH4 time series over Darwin where low values of XCH4 are observed for several days.

Figure 100: Map showing XCH4 concentrations above Greenland and parts of North America for three years averaged data with pixels with qa_value > 0.5. The XCH4 pixels in Greenland shows outliers along the mountain and coastline regions. Similar features are observed at the Antarctic.
11.4 First assessment of L2_CH4 OFFL processor update to version 02.02.00

On July 1, 2021, the L2_CH4 OFFL processor was updated to version 02.02.00 with different spectroscopic parameters (see the S5P CH4 Product Readme File). This section provides a first assessment of the data accuracy for processor version 02.02.00. From Table 14 the impact of the update in spectroscopic parameters is observed in a small change in the relative bias: from 0.4% for processor 01.xx.xx prior to July 2021 to 0.2% for processor 02.xx.xx. The latter estimate should be considered as preliminary and must be confirmed with consolidated NDACC data and when a longer time series is available, and with recent TCCON data.

<table>
<thead>
<tr>
<th>station</th>
<th>rel. std</th>
<th>correla</th>
<th>rel diff bias(%)</th>
<th>std rel diff (%)</th>
<th>#</th>
<th>rel. std</th>
<th>correla</th>
<th>rel diff bias(%)</th>
<th>std rel diff (%)</th>
<th>lat</th>
</tr>
</thead>
<tbody>
<tr>
<td>THULE</td>
<td>58</td>
<td>0.8</td>
<td>0.8</td>
<td>0.91</td>
<td>58</td>
<td>0.8</td>
<td>0.8</td>
<td>1.2</td>
<td>0.94</td>
<td>76.5</td>
</tr>
<tr>
<td>KIRUNA</td>
<td>60</td>
<td>0.7</td>
<td>0.79</td>
<td>-1.55</td>
<td>0.5</td>
<td>60</td>
<td>0.7</td>
<td>-3.08</td>
<td>0.54</td>
<td>67.8</td>
</tr>
<tr>
<td>SODANKYLA</td>
<td>114</td>
<td>0.7</td>
<td>0.16</td>
<td>-0.95</td>
<td>1.08</td>
<td>114</td>
<td>0.6</td>
<td>-2.92</td>
<td>1.22</td>
<td>67.4</td>
</tr>
<tr>
<td>ST.PETERSBURG</td>
<td>14</td>
<td>0.2</td>
<td>0.07</td>
<td>-0.28</td>
<td>0.96</td>
<td>14</td>
<td>0.1</td>
<td>-1.9</td>
<td>1.03</td>
<td>59.9</td>
</tr>
<tr>
<td>GARMISCH</td>
<td>271</td>
<td>0.6</td>
<td>0.36</td>
<td>-0.05</td>
<td>0.78</td>
<td>271</td>
<td>0.6</td>
<td>-1.19</td>
<td>0.88</td>
<td>47.5</td>
</tr>
<tr>
<td>ZUGSPITZEN</td>
<td>111</td>
<td>1.6</td>
<td>-0.12</td>
<td>-0.4</td>
<td>1.19</td>
<td>111</td>
<td>1.6</td>
<td>-1.8</td>
<td>1.05</td>
<td>47.4</td>
</tr>
<tr>
<td>JUNGFRAUJOCH</td>
<td>177</td>
<td>1.7</td>
<td>0.48</td>
<td>0.38</td>
<td>1.36</td>
<td>177</td>
<td>1.9</td>
<td>-0.42</td>
<td>1.34</td>
<td>46.5</td>
</tr>
<tr>
<td>TORONTO</td>
<td>159</td>
<td>2.4</td>
<td>0.35</td>
<td>2.14</td>
<td>2.32</td>
<td>159</td>
<td>2.1</td>
<td>0.38</td>
<td>1.1</td>
<td>43.6</td>
</tr>
<tr>
<td>BOULDER.CO</td>
<td>1126</td>
<td>1.5</td>
<td>0.62</td>
<td>-0.65</td>
<td>0.82</td>
<td>1126</td>
<td>1.4</td>
<td>0.71</td>
<td>-0.96</td>
<td>0.73</td>
</tr>
<tr>
<td>MAUNA.LOA.HI</td>
<td>24</td>
<td>1.9</td>
<td>0.1</td>
<td>1.11</td>
<td>2.01</td>
<td>24</td>
<td>2.5</td>
<td>-0.05</td>
<td>1.91</td>
<td>19.5</td>
</tr>
<tr>
<td>LA.REUNION.MAIDO</td>
<td>8</td>
<td>3</td>
<td>0.73</td>
<td>-0.09</td>
<td>0.63</td>
<td>8</td>
<td>2.7</td>
<td>-0.51</td>
<td>0.67</td>
<td>-21.1</td>
</tr>
<tr>
<td>WOLLONGONG</td>
<td>263</td>
<td>1.2</td>
<td>0.23</td>
<td>-0.45</td>
<td>1.29</td>
<td>263</td>
<td>1.2</td>
<td>-1.49</td>
<td>1.33</td>
<td>-34.4</td>
</tr>
<tr>
<td>LAUDER</td>
<td>490</td>
<td>1.6</td>
<td>0.33</td>
<td>-0.5</td>
<td>1</td>
<td>490</td>
<td>1.7</td>
<td>-0.97</td>
<td>0.97</td>
<td>-45.0</td>
</tr>
<tr>
<td>ARRIVALHEIGHTS</td>
<td>58</td>
<td>0.4</td>
<td>0.42</td>
<td>0.96</td>
<td>1</td>
<td>58</td>
<td>0.4</td>
<td>-1.03</td>
<td>0.87</td>
<td>-77.8</td>
</tr>
</tbody>
</table>

Table 14 - NDACC network statistics comparing the OFFL data with processor version 02.xx.xx (since July 1 2021) for both the bias corrected and standard product. The update in July introduced new spectroscopic parameters which have a systematic effect and this is seen in the bias estimate. The estimate uses the rapid delivery NDACC data and should be consolidated in future reports.

11.5 Validation of L2_CH4 OFFL sun-glint retrievals

Since the update to processor 02.03.01 on November 14 2021, pixels over the ocean are available. Ocean measurements performed over sun-glint geometry improve significantly the TROPOMI XCH4 product coverage. Since sun-glint geometry depends on the position of the sun relative to the satellite, these measurements have a seasonal cycle and different parts of the ocean are covered by the sun-glint geometry at different times of the year. It covers approximately an area of 30 degrees extension in latitude (see Figure 101 – right). In November-February it covers mostly the Northern hemisphere, and April-September it covers mostly the Southern hemisphere. The typical annual coverage is shown in Figure 101 – left, which corresponds to the year 2020.
To evaluate the sun-glint retrievals introduced in the processor 02.03 in November 2021, the scientific L2 product available at SRON for the period between November 2017 and September 2021 is used. A similar validation approach is used as for the standard methane L2 data, except that an additional filtering is enabled to keep only sun-glint pixels over ocean. Only the bias-corrected methane values in the scientific L2 products are considered.

The systematic difference between the S5P scientific L2.CH4 bias-corrected product and the TCCON XCH4 data is on average -0.55%, well within the mission requirements. The 1σ spread of the relative difference between the S5P and the TCCON data around the mean value is 0.7%, also below the mission requirements (precision <1%). The individual values for the different stations are indicated in Figure 103.
Figure 103: Chart of relative mean difference between S5P scientific L2_CH4 and FTIR CH4 column data at 11 TCCON stations where colocation were found with the sun-glint pixels. The stations are sorted with decreasing latitude. The relative mean difference of the xCH4 exceeds the mission requirements for JPL site where only two days of co-location were found.

For NDACC 6 sites have co-locations with the sun-glint pixels. These are a mixture of island sites or sites close to a coastline. Although the number of co-locations is low, the bias and precision are close to the values of the standard bias-corrected.

<table>
<thead>
<tr>
<th>station</th>
<th>#</th>
<th>rel. std</th>
<th>correl.</th>
<th>rel. diff</th>
<th>bias(%)</th>
<th>std rel diff (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TORONTO</td>
<td>47</td>
<td>4.4</td>
<td>-0.1</td>
<td>1.28</td>
<td>3.91</td>
<td></td>
</tr>
<tr>
<td>IZANA</td>
<td>98</td>
<td>0.8</td>
<td>0.74</td>
<td>-1.56</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>MAUNA.LOA.HI</td>
<td>18</td>
<td>1.9</td>
<td>0.43</td>
<td>-0.17</td>
<td>1.42</td>
<td></td>
</tr>
<tr>
<td>PARAMARIBO</td>
<td>5</td>
<td>0.9</td>
<td>0.21</td>
<td>-2.53</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>REUNION.MAIDO</td>
<td>38</td>
<td>1.9</td>
<td>0.03</td>
<td>-1.02</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>WOLLONGONG</td>
<td>14</td>
<td>1.1</td>
<td>-0.42</td>
<td>-1.23</td>
<td>1.36</td>
<td></td>
</tr>
</tbody>
</table>

Table 15 – Overview of statistical quality indicators for the co-located S5P sun-glint pixels and NDACC. Bias and dispersion are of the same order of magnitude as the values for the standard S5P bias corrected product. Correlations should be ignored due to the low number of co-locations.
12 Validation Results: L2_CLOUD

12.1 L2_CLOUD products and requirements

This section reports on the validation of the following geophysical variables of the S5P TROPOMI L2_CLOUD product identified in Table 1: the (radiometric) Cloud Fraction (CF), the Cloud Top Height (CTH)/Cloud Height (CH), and the Cloud Optical Thickness (COT)/Cloud Albedo (CA). There are actually two sub-products within the L2_CLOUD files: OCRA/ROCINN_CAL (providing CF, CTH, COT) and OCRA/ROCINN_CRB (providing CF, CH, CA). Shorthand notation used here for both sub-products is CLOUD CAL and CLOUD CRB. Validation results are discussed with respect to the product quality targets outlined in Table 3. Before the UPAS processor switch to version 2.1.3 in July 2020, the NRTI and OFFL processors are based on the same algorithm and produce very similar data products; the situation is different after the switch, the NRTI and OFFL cloud data processors differing. Therefore, Subsection 12.3 concentrates on the validation of the L2_CLOUD OFFL product while Subsection 12.4 demonstrates evidence that NRTI and OFFL data do not differ significantly and that their respective validations yield similar conclusions before the processor switch. The latest upgrade to version 02.03.00, issued in March 2022, will be assessed in a next update of this report.

12.2 Validation approach

12.2.1 Ground-based networks

CLOUDNET lidar/radar data

S5P TROPOMI L2_CLOUD cloud data have been routinely compared at 23 ground-based stations (Table 16) to reference lidar/radar data from the cloud target classification product of the CLOUDNET and ARM ground-based networks [ER_Cloudnet]. Cloud base height, cloud top height and a vertical cloud classification profile (resolution <100 m) are provided each 30 s, typically.

Comparison settings

For the comparisons between S5P and CLOUDNET data, two approaches were tested.

First approach: S5P TROPOMI pixels are selected if qa_value > 0.5, cloud_fraction > 0.5, the pixel encompasses the CLOUDNET station, and the cloud is not multi-layered according to the CLOUDNET classification. Per S5P overpass, the closest co-location pair in time (within a time interval of 600 s) only is kept. This approach was routinely used in ROCVR validation reports up to ROCVR update #03, but the constraints to high cloud fractions and monolayer cloud limited the scope of the validation.

Second approach: S5P TROPOMI pixels are selected if qa_value > 0.5, cloud_fraction > 0.1, the pixel encompasses the CLOUDNET station, the station is cloud covered (according to CLOUDNET) at least 50% of the 1200 s temporal interval centered at the TROPOMI overpass time, and the standard deviation of CLOUDNET cloud height is smaller than 0.5 km within this temporal interval. Note that there is no filtering of multilayer clouds. The average cloud height or cloud top height is calculated from CLOUDNET cloud types 1-7. Although with this second approach generally a higher bias is obtained, correlative properties also improve or are comparable. Given the broader scope of this second approach, it is selected here.

We present here also comparisons of the S5P TROPOMI FRESCO, which employs an alternative cloud retrieval algorithm, with CLOUDNET, using the same comparison settings.
<table>
<thead>
<tr>
<th>Station</th>
<th>Location</th>
<th>Network</th>
<th>Latitude (N)</th>
<th>Longitude (E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ny-Ålesund</td>
<td>Svalbard</td>
<td>CLOUDNET</td>
<td>78.932</td>
<td>11.921</td>
</tr>
<tr>
<td>Summit</td>
<td>Greenland</td>
<td>NOAA/ARM</td>
<td>72.60</td>
<td>-38.42</td>
</tr>
<tr>
<td>Andoya</td>
<td>Norway</td>
<td>ARM</td>
<td>69.14</td>
<td>15.68</td>
</tr>
<tr>
<td>Kenttarova</td>
<td>Finland</td>
<td>CLOUDNET</td>
<td>67.987</td>
<td>24.23</td>
</tr>
<tr>
<td>Hyttyala</td>
<td>Finland</td>
<td>CLOUDNET</td>
<td>61.84</td>
<td>24.29</td>
</tr>
<tr>
<td>Norunda</td>
<td>Sweden</td>
<td>CLOUDNET</td>
<td>60.09</td>
<td>17.48</td>
</tr>
<tr>
<td>Mace Head</td>
<td>Ireland</td>
<td>CLOUDNET</td>
<td>53.325</td>
<td>-9.9</td>
</tr>
<tr>
<td>Lindenberg</td>
<td>Germany</td>
<td>CLOUDNET</td>
<td>52.211</td>
<td>14.13</td>
</tr>
<tr>
<td>Leipzig</td>
<td>Germany</td>
<td>CLOUDNET</td>
<td>51.35</td>
<td>12.43</td>
</tr>
<tr>
<td>Chilbolton</td>
<td>United Kingdom</td>
<td>CLOUDNET</td>
<td>51.145</td>
<td>-1.437</td>
</tr>
<tr>
<td>Jülich</td>
<td>Germany</td>
<td>CLOUDNET</td>
<td>50.909</td>
<td>6.413</td>
</tr>
<tr>
<td>Palaiseau</td>
<td>France</td>
<td>CLOUDNET</td>
<td>48.713</td>
<td>2.208</td>
</tr>
<tr>
<td>Munich</td>
<td>Germany</td>
<td>CLOUDNET</td>
<td>48.15</td>
<td>11.57</td>
</tr>
<tr>
<td>Schneefernerhaus</td>
<td>Germany</td>
<td>CLOUDNET</td>
<td>47.42</td>
<td>10.98</td>
</tr>
<tr>
<td>Bucharest</td>
<td>Romania</td>
<td>CLOUDNET</td>
<td>44.35</td>
<td>26.03</td>
</tr>
<tr>
<td>Potenza</td>
<td>Italy</td>
<td>CLOUDNET</td>
<td>40.6</td>
<td>15.72</td>
</tr>
<tr>
<td>Graciosa</td>
<td>Azores</td>
<td>ARM</td>
<td>39.092</td>
<td>-28.026</td>
</tr>
<tr>
<td>Granada</td>
<td>Spain</td>
<td>CLOUDNET</td>
<td>37.16</td>
<td>-3.605</td>
</tr>
<tr>
<td>Mindelo</td>
<td>Cape Verde</td>
<td>CLOUDNET</td>
<td>16.879</td>
<td>-24.995</td>
</tr>
<tr>
<td>Iquique</td>
<td>Chile</td>
<td>CLOUDNET</td>
<td>-20.54</td>
<td>-70.18</td>
</tr>
<tr>
<td>Maito</td>
<td>La Reunion</td>
<td>CLOUDNET</td>
<td>-21.08</td>
<td>55.38</td>
</tr>
<tr>
<td>Villa Yacanto</td>
<td>Argentina</td>
<td>ARM</td>
<td>-32.13</td>
<td>-64.73</td>
</tr>
<tr>
<td>Punta-Arenas</td>
<td>Chile</td>
<td>CLOUDNET</td>
<td>-53.135</td>
<td>-70.883</td>
</tr>
</tbody>
</table>

Table 16 – List of ground-based stations providing the cloud classification data product, and used in this study: 19 CLOUDNET stations and 4 ARM (Atmospheric Radiation Measurement) stations. Data is collected through EVDC.

### 12.2.2 Satellites

**MODIS, NPP VIIRS and OMI O₃-O₂**

S5P TROPOMI L2_CLOUD cloud data (internal UPAS product, comparable to the operational OFFL 01.01.05 product) have also been compared to MODIS L3 data (https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MYD08_D3/) and the VIIRS non-operational product¹ made available by NASA in 2018. The comparison with MODIS allows only for daily means validation whereas the comparison against VIIRS offers a pixel-by-pixel validation of the product.

¹ The VIIRS cloud datasets were obtained from a pre-production code run specifically for limited S5P team analysis. The VIIRS cloud algorithm is based on the MODIS Collection 6 algorithms [https://modis-atmosphere.gsfc.nasa.gov/documentation/collection-6; Platnick et al. (2017). The CLDPROP data were released in Feb. 2019 and are described here: https://modis-
Zonal means of cloud fraction and cloud (top) height of OMI O2-O2 (OMCLDO2), MODIS and S5P FRESCO are compared to those of S5P L2_CLOUD CAL and CRB. Note that here, radiometric cloud fractions are scaled to a cloud albedo of 0.8. Regarding cloud (top) height comparisons, cloud height with scaled radiometric cloud fraction < 0.05 are removed (does not apply to MODIS).

**Comparison settings**

For the comparisons between S5P L2_CLOUD and VIIRS data, the following exclusion filters were applied: TROPOMI with qa_value < 0.5 were rejected; snow/ice scenes as well; VIIRS geometrical cloud fraction < 0.9 (to mitigate regridding artefacts); CTH > 15 km (as the S5P L2_CLOUD algorithm does not retrieve above this value); COT < 1 (as the S5P L2_CLOUD algorithm does not retrieve below this value), and COT > 150 (as this is the maximum VIIRS value after regridding).

12.2.3 Alternative S5P cloud algorithms

**S5P FRESCO**

The support product S5P TROPOMI FRESCO cloud height is also compared to CLOUDNET observations, and directly with CLOUD CRB at the CLOUDNET stations. This helps to judge if discrepancies between S5P CLOUD CRB and CLOUDNET are specific to the adopted cloud retrieval algorithm or are of more general nature. The S5P FRESCO support product is not publicly disseminated separately, but is used as input for e.g., the S5P NO2 retrieval. Earlier versions of the algorithm are described in e.g., [Koelemeijer 2001]. Like CLOUD CRB, FRESCO-S models a cloud as a Lambertian reflector. Information on cloud pressure and effective cloud fraction is derived from the reflectance in and around the O2 A band. As opposed to CLOUD CRB, where cloud albedo is retrieved, in FRESCO-S, the cloud albedo is assumed to be 0.8 or the TOA reflectance at 758 nm if the reflectance is larger than 0.8. We note that at small cloud fractions, the surface albedo is adapted to prevent negative cloud fractions.

**Comparison settings**

Given the different assumption for cloud albedo in the CLOUD CRB and FRESCO retrieval models, CLOUD CRB CF and FRESCO CF are not directly comparable. Instead, we compare here the cloud fractions rescaled to cloud albedo=0.8: CF_rescaled = CF*CA/0.8. Also called sRCF (scaled radiometric cloud fraction) here.

S5P CLOUD pixels and S5P FRESCO pixels covering CLOUDNET stations were extracted. Common overpasses were considered. For CLOUD, qa_value ≥ 0.25 or ≥ 0.5 was applied; both settings are discussed in the Readme file. For FRESCO, qa_value > 0.5 was applied. Additionally, for the cloud height comparisons, CF_rescaled > 0.1 was applied

12.2.4 Field campaigns and modelling support

None for this report.

---

Those operational publicly available data might have some differences compared to the limited data provided by the NASA group directly to DLR in 2018.
12.3 Validation of L2_CLOUD OFFL

12.3.1 Recommendations for data usage followed

As recommended, only those TROPOMI ground pixels associated with a qa_value above 0.5 have been assessed here. The qa_value summarizes the quality of the product by taking into consideration several aspects like the spectral channel quality flags from L1B data, geometry limitations (e.g. not reliable retrievals for SZA>75°), inhomogeneous scene warnings, high residual of the fitting process etc.

Some of the known data quality issues are not covered by the quality flags and have been considered when interpreting the validation results reported hereafter (see also the Product Readme File (PRF)). Those issues are:

1. instrumental feature: spatial mis-registration between TROPOMI bands 3-4 (OCRA, UV trace gas fitting window) and band 6 (ROCINN fitting window),
2. insensitivity to very thin clouds,
3. treatment of multi-layer clouds,
4. treatment of ice clouds,
5. snow/ice conditions,
6. unknown straylight impact in the NIR,
7. saturation (note that the L1B flagging works well, only blooming isn’t flagged correctly yet),

For further details, data users are encouraged to read the Product Readme File (PRF), Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) associated with this data product, all on https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms.

12.3.2 Status of validation

This section presents a summary of the key validation results obtained by the MPC VDAF and by S5PVT AO projects. It is based on regular updates of the results reported at S5PVT workshops. Up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal at http://mpc-vedaf.tropomi.eu.

The validation vs. CLOUDNET ground-based data uses S5P L2_CLOUD RPRO+OFFL 01.01.07-02.03.00 data. This covers the time period from 2018-04-30 till June 2022. CLOUDNET data from 23 stations were considered in this analysis (two stations added since ROCVR#14). Note that the station Kenttarova did not provide valid co-locations in the CLOUDNET measurement period (20191027-20191225). The format change issue at the CLOUDNET database is resolved, so plots involving CLOUDNET data were updated.

12.3.3 Radiometric cloud fraction (L2_CLOUD CAL & L2_CLOUD CRB)

12.3.3.1 Bias

Comparison with other satellites and alternative algorithm S5P FRESCO

In Compernolle et al., (2021), the latitudinal variation of zonal means of CLOUD CAL and CRB v1 cloud fraction is also compared with MODIS geometric cloud fraction. A similar variation is found between the products, but as expected, the geometric cloud fraction is higher than the radiometric cloud fraction of the SSP CLOUD products.
Figure 104. Latitudinal variation of zonal means of scaled radiometric cloud fraction of S5P CLOUD OCRA/ROCINN_CRB and S5P FRESCO at 4 days, with different product versions. 2020/07/09: ROCINN_CRB v1, FRESCO v1.3. 2020/07/18: ROCINN_CRB v2.1, FRESCO v1.3. 2020/11/30: ROCINN_CRB v2.1, FRESCO v1.4. 2021/07/02: ROCINN_CRB v2.2, FRESCO v2.2. Selection criteria: qa_value>0.5, no invalid values. ‘common’ means the subset of valid pixels common to CLOUD and FRESCO.
Figure 105. Similar as Figure 104, but for a recent day, where CLOUD CRB is at version 2.2.1 and FRESCO at 2.3.1. Two quality filter criteria are used for CRB: the standard qa_value>0.5, and a looser qa_value>0.25.

Figure 104 and Figure 105 presents the latitudinal variation of zonal means of scaled radiometric cloud fraction of S5P CLOUD OCRA/ROCINN_CRB and S5P FRESCO at 5 days with different product versions.

- At 2020/07/09, with ROCINN_CRB v1, FRESCO v1.3. S5P FRESCO sRCF is larger than S5P CLOUD ROCINN_CRB sRCF. Subsetting to the set of common pixels has no large impact.

- For the other cases, (i.e., with ROCINN_CRB v2), (i) the impact of taking the set of common pixels is larger, as qa_value>0.5 becomes stricter for ROCINN, and (ii) the difference between both product sRCFs becomes smaller. This stays true for the recent day 2021/11/29, where the version for ROCINN is 2.2.1 and for FRESCO is 2.3.1. The excess FRESCO pixels have a low cloud fraction.

- The recommendation of qa_value>0.5 for ROCINN_CRB is only a guideline and is too strict for some applications. When using a more loose qa_value>0.25 as proposed in the CLOUD PRF (Figure 105), the behaviour is similar (CRB sRCF slightly lower than FRESCO sRCF), indicating that this less strict quality filtering can indeed be used.

Figure 106 presents boxplots of the difference (top left) and normed relative difference (top right) between rescaled cloud fractions of CLOUD CRB and FRESCO at the CLOUDNET stations, as well as the evolution of the monthly median difference over the CLOUDNET stations (bottom). The loose criterion qa_value>0.25 was applied, but this has only a small impact on the resulting figures compared to the qa_value>0.5 criterion. Overall, mean and median difference are negative (CRB lower than FRESCO) but in most cases within the 20% bias requirement. The strong deviations at the Northern stations Kenttarova, Andoya, Summit and Ny-Ålesund are at least partly due to pixels with snow-ice cover. Note that also at Maido (Reunion) a stronger deviation is seen. Since April 2021, more months with CRB sRCF > FRESCO occur, especially at Munich.

There is an important impact of the version upgrade from CLOUD v1 to v2.1 (2020/07/13): the median difference becomes lower for most stations. In particular, the strongly negative difference at the Northern station Ny-Ålesund becomes positive. This can at least partly be attributed to the better surface albedo treatment of this complicated sea/land/snow/ice station, and to a more strict snow/ice filtering. Strong reductions at the other high-latitude sites Andoya, Kenttarova, and island sites Maido and Mindelo are also visible. At the site Punta-Arenas the negative bias becomes slightly positive. The FRESCO upgrade from
v1.3 to v1.4 (2020/11/29), the upgrade to CLOUD v2.2 and FRESCO v2.2 (both at 2021/07/01) and the upgrade to FRESCO v2.3 (2021/11/14) have no important impact on the sRCF bias.

Figure 106. Top. Box plots of S5P CLOUD CRB OFFL cloud fraction minus S5P FRESCO OFFL cloud fraction, after both have been rescaled to a cloud albedo of 0.8. Middle: same but for the normed relative difference. Bottom: Monthly median of scaled cloud fraction difference, per month and per station. Indicated are the pixel size switch (2019/08/06), the CLOUD v1 to v2.1 upgrade (2020/07/13), the FRESCO v1.3 to 1.4 upgrade (2020/11/29), the upgrade from CLOUD v2.1 to v2.2 and from FRESCO v1.4 to v2.2 (both on 2021/07/01), the upgrade from FRESCO v2.2 to v2.3 (2021/11/14) and the upgrade from CLOUD v2.2 to v2.3 (2022/03/06).
12.3.3.2 Dispersion

**Figure 107** (left) presents the evolution of the difference dispersion (½ IP68 (CRB minus FRESCO)) over the CLOUDNET stations, per month. **Figure 107** (right) presents the overall dispersion per station. At most stations, the ½ IP68 slightly exceeds the dispersion requirement of 0.05, while there is a high dispersion at the high-latitude stations Kenttarova, Andoya, Summit and Ny-Ålesund. This is likely due to a different treatment of snow/ice by both algorithms (see also section 12.3.3.3). Note also the high dispersion at Maida (Reunion Island). The dispersion at Ny-Ålesund,Andoya, Kenttarova and Maida is reduced with the upgrade from CLOUD v1 to CLOUD v2.1 (2020/07/13). Note that after this upgrade, there is no more data from Summit, due to a stricter snow/ice filtering.

**Figure 107.** Left: Monthly ½ IP68 of scaled cloud fraction difference of CLOUD OFFL vs FRESCO OFFL, per month and per station. Indicated are the pixel size switch (2019/08/06), the CLOUD v1 to v2 upgrade (2020/07/13), the FRESCO v1.3 to 1.4 upgrade (2020/11/29), the upgrade from CLOUD v2.1 to v2.2 and from FRESCO v1.4 to v2.2 (both on 2021/07/01), the upgrade from FRESCO v2.2 to v2.3 (2021/11/14) and the upgrade from CLOUD v2.2 to v2.3 (2022/03/06). Right: ½ IP68 of scaled cloud fraction difference (CLOUD CRB-FRESCO), per station. Note that the dispersion at the station Summit is off-scale and extends to ~0.6.

12.3.3.3 Dependence on influence quantities

The S5P L2_CLOUD cloud fraction gets unphysically high values at very large SZAs (above 85 degrees) due to very weak illumination. The other cloud parameters might also be affected for high SZAs due to limitation in the RTM treatment of spherical atmosphere. The high surface albedo above snow and/or ice covered surfaces is a challenge for cloud retrievals. Note that a very large SZA implies a measurement above the polar region, and therefore snow-ice covered surfaces are likely.

12.3.3.4 Drifts, cycles and shorter term variability

The presence of degradation in the L1b version 1 data may lead to a degradation particularly in the OCRA cloud\_fraction\_apriori and also to a lesser degree in the ROCINN parameters. This effect should be minimized with the L1b data version 2 update, which contains a degradation correction and was introduced very recently in UPAS 2.2.1. More measurements and reprocessing of existing data will be needed to verify this.

12.3.3.5 Geographical patterns

The effects of the solar zenith angle and surface albedo mentioned above give rise to geographical patterns. Furthermore, cloud parameters in UPAS 1.1.x may show an enhancement at the east edge of the swath for some months at certain latitudes. The effect seems to be strongest in the latitude bands [40-60]N and [30-40]S. This issue is reduced in UPAS 2.1.3 operational since July 2020. An example for
cloud_fraction_apriori (the cloud fraction as determined by OCRA; note that due to strong regularization, the retrieved CAL and CRB cloud fractions are close to this one) is shown in Compernolle et al. (2021), Figure S28.

12.3.4 Cloud top height (L2_CLOUD CAL) and cloud height (L2_CLOUD CRB)

12.3.4.1 Bias

Comparison with other satellite and alternative S5P cloud height retrievals

In Compernolle et al. (2021), S5P CLOUD v1 CAL and CRB cloud heights are also compared to MODIS and VIIRS CTH. Conclusions are as follows: (i) CAL CTH is on average below VIIRS CTH, but it matches well the low CTH mode typical for liquid clouds, (ii) a similar latitudinal variation with MODIS CTH and OMCLDO2 CH is found, but with offsets.

Figure 108: Latitudinal variation of zonal means of cloud height of S5P CLOUD OCRA/ROCINN_CRB and S5P FRESCO at 4 days, with different product versions. 2020/07/09: ROCINN v1, FRESCO v1.3. 2020/07/18: ROCINN v1.3, 2020/11/30: ROCINN v2.1, FRESCO v1.4. 2021/07/02: ROCINN v2.2, FRESCO v2.2. 2022/02/27: ROCINN: v2.2, FRESCO v2.3. Selection criteria: qa_value>0.5, no invalid values, sRCF>0.05. ‘common’ means the subset of valid pixels common to CLOUD and FRESCO. For2022/02/27, the ROCINN_CAL cloud top height is added to the plot.
Figure 108 presents the latitudinal variation of zonal means of cloud height of S5P CLOUD OCRA/ROCINN_CRB and S5P FRESCO at 4 days with different product versions.

- At 2020/07/09, with ROCINN_CRB v1, FRESCO v1.3, the mean FRESCO cloud height is considerably lower than that of ROCINN_CRB. However, when taking the common subset of pixels, both cloud heights become comparable. This is probably due to FRESCO pixels with low cloud fraction and low cloud height that are removed in the common subset selection.

- At 2020/07/18, with ROCINN_CRB v2.1 and FRESCO v1.3, for the common subset, CLOUD CRB cloud height becomes slightly lower than that of FRESCO.

- At 2020/11/30, with ROCINN_CRB v2.1 and FRESCO v1.4, for the common subset, FRESCO cloud height becomes considerably higher than that of ROCINN_CRB. This stays true at 2021/07/02, with ROCINN_CRB v2.2 and FRESCO v2.2, and 2022/02/27, with ROCINN_CRB v2.2 and FRESCO v2.3. For this day, we also investigated the less strict CRB quality filter qa_value>0.25, but this did not change the overall outcome (not shown). Note that FRESCO cloud height is in between ROCINN_CRB cloud height and ROCINN_CAL cloud top height.

S5P CLOUD CRB CH was compared with S5P FRESCO CH, over the CLOUDNET stations (Figure 109). At most stations, a small bias is encountered between 0 and -0.5 km (mean and median difference; CRB lower than FRESCO) or ~between -10% and -20% (median normed relative difference). This is within the 20% bias requirement. Clear exceptions (CRB higher than FRESCO) occur at the stations Iquique, Ny-Ålesund and Summit. This boxplot however hides the important impact of version upgrades (Figure 109, bottom left). Initially, the monthly median difference is small in most cases (between 0 and -0.2 km). With the upgrade from CLOUD v1 to v2, the monthly median difference becomes more negative (CRB lower than FRESCO) but mostly still within 0.5 km. The positive bias at Ny-Ålesund becomes negative, while the positive biases at Mindelo and Iquique mostly disappear. The issue of exceptionally high CRB values is resolved with the upgrade to CLOUD v2.1. At 2020/11/29, the FRESCO OFFL v1.3 to v1.4 was introduced. This results in an even larger negative difference (CRB lower than FRESCO) which can reach over 1 km. The upgrades to CLOUD v2.2 and FRESCO v2.2 (both on 2021/07/01) do not seem to have a strong impact on their median difference cloud height; at most a small reduction. The upgrade to FRESCO 2.3 (2021/11/14) also does not seem to have an important impact.

We also compare S5P CLOUD CAL CTH with S5P FRESCO CH. The high impact of the FRESCO upgrade from v1.3 to v1.4 (2020/11/29) is also here visible. Before, CAL CTH is about 1 km above FRESCO CH, while after the upgrade both are more comparable. Furthermore, there is a bias reduction at Mindelo with the upgrade to CLOUD v2.2 and FRESCO v2.2 (2021/07/01).

Comparison with CLOUDNET cloud top height and cloud height

The previous paragraphs made clear that the CLOUD (v1 to v2) and FRESCO version (v1.3 to 1.4) upgrades have an important impact on their intercomparison. Here, we check how the effective cloud heights of CLOUD CAL, CLOUD CRB and FRESCO are related to the cloud heights as obtained from CLOUDNET, before and after their version upgrades.

Figure 110 to Figure 112 compare average cloud height values of S5P CLOUD CAL, S5P CLOUD CRB and S5P FRESCO with cloud heights (base, mean, and top) of CLOUDNET for CLOUD v1, CLOUD v2, FRESCO v1.3 and FRESCO v1.4, separately for low and high clouds (separation at CLOUDNET CTH=4km).
Conclusions are as follows:

- **CLOUD CAL v1 CTH** is slightly above the CLOUDNET cloud mean height (CMH) for high clouds and **CLOUD CAL v2 CTH** is slightly below. For low clouds it is mostly in between CLOUDNET CTH and CMH.

- **CLOUD CRB v1 and v2 CH** are in between CLOUDNET CMH and CLOUDNET CBH for high clouds, but v2 is slightly lower than v1. For low clouds CLOUD CRB v1 is close to the CLOUDNET CBH, and even below for CLOUD CRB v2.

![Figure 109](image)

Figure 109: Top. Boxplots of S5P CLOUD CRB CH minus S5P FRESCO CH (left) and of the normed difference (right). Bottom. Monthly median difference per station of CRB - FRESCO (left) and CAL - FRESCO (right). Indicated are the pixel size switch (2019/08/06), the CLOUD v1 to v2 upgrade (2020/07/13), the FRESCO v1.3 to 1.4 upgrade (2020/11/29), the upgrade from CLOUD v2.1 to v2.2 and from FRESCO v1.4 to v2.2 (both on 2021/07/01), the upgrade to FRESCO v2.3 (2021/11/14) and the upgrade from CLOUD v2.2 to v2.3 (2022/03/06).

- The exceptional overshoot of CLOUD cloud height value at Ny-Ålesund disappears with the upgrade from CLOUD v1 to CLOUD v2.

- For high clouds, FRESCO CH is slightly below CLOUDNET CMH, before and after the upgrade to 1.4. The upgrade has a bigger impact for low clouds: while before the upgrade, it is close to CLOUDNET CBH and CMH, after the upgrade it moves close to CLOUDNET CTH. For the Northern stations Norunda, Hyytiala and Ny-Ålesund, there is even an overshoot. Further analysis reveals this seems related to a seasonal feature rather than a version change. S5P FRESCO CH at high-latitude (Norunda, Hyytiala, Ny-Ålesund) stations present a seasonal cycle in the bias, with high FRESCO CH early 2020 and 2021 (see Figure 113). Note that S5P CLOUD CRB CH data is mostly filtered out in this time period and place. This is likely related to snow-ice and/or high SZA conditions.
As for CLOUD CAL, the parameter is retrieved as a *cloud top height*, while for CRB and FRESCO they are retrieved as *cloud heights*, we compare in what follows CLOUD CAL CTH with CLOUDNET CTH, CLOUD CRB CH with CLOUDNET CMH and FRESCO CH with CLOUDNET CMH.

**Figure 110.** Mean values of the offset from the CLOUDNET sensor altitude, per station of (i) S5P CLOUD OFFL CAL cloud top height, (ii) CLOUDNET cloud base height, (iii) CLOUDNET cloud mean height (calculated from cloudy altitude bins), (iv) CLOUDNET cloud top height. Top/bottom row panels: high and low clouds respectively (split at CLOUDNET CTH=4km). Left/right column panels: CLOUD v1 and CLOUD v2 respectively. Only stations with data for both versions are included. Data count per station is given on the plot.
Figure 111. Same as Figure 110 but for S5P CLOUD OFFL CRB CH instead of S5P CLOUD OFFL CAL CTH.

Figure 112. Same as Figure 110 but for S5P OFFL FRESCO CH instead of S5P CLOUD OFFL CAL CTH, and the version split at FRESCO v1.3 and FRESCO ≥v1.4.
Figure 113. Left. Time series of S5P CLOUD CRB OFFL CH and CLOUDNET CMH. Right. Time series of S5P FRESCO OFFL CH and CLOUDNET CMH. Note the peak in FRESCO CH in early 2020 and 2021, while CLOUD CRB has almost no data in this period.

L2_CLOUD CAL cloud top height is generally below the CLOUDNET cloud top height. This can be seen in Figure 114, which presents boxplots per station of absolute-scale and relative difference (top panels) and of the monthly median difference (bottom panel). The initially positive bias at Ny-Ålesund becomes negative with the upgrade to CLOUD v2. High negative biases at Granada (exceeding 3 km) and a more negative bias at Jülich are obtained after the UPAS upgrade. We note furthermore that the relatively higher negative biases at Lindenberg and Munich are probably related to the higher proportion of high-lying ice clouds (Compernolle et al., 2021). The CLOUD CAL algorithm works better for liquid clouds.

The bias depends on the cloud height. In distribution plots of CLOUDNET cloud top heights two modes are typically visible (see e.g., Compernolle, 2021, Fig. 10), where the higher mode contains more ice cloud and multilayer clouds. We consider a CLOUDNET CTH of 4 km as the threshold between low (<4 km) and high (>4 km) clouds. To obtain an overall value for bias and relative bias, we calculate the median over all per-station medians of the difference and of the relative difference, for low and high clouds, and for CLOUD version 1 and version 2. The bias (relative bias) is -2 km (-30%) for high clouds for v1, -3 km (-40%) for v2, and -0.3 km (-15%) for low clouds (both v1 and v2) (see Table 17). Note that for v2, the number of stations is more limited and co-locations at some stations are sparse.
Table 17. Statistics, calculated over the stations of S5P CLOUD OFFL CAL CTH-CLOUDNET CTH, S5P CLOUD OFFL CRB CH-CLOUDNET CMH and S5P FRESCO OFFL CH-CLOUDNET CMH. The Pearson-R is calculated over the individual measurement pairs comparing with both CLOUDNET CMH and CLOUDNET CTH. Measurements up to 2022/06/02 are included.

<table>
<thead>
<tr>
<th>CTH: CAL - CLOUDNET</th>
<th>#stations</th>
<th>Med(meddiff)</th>
<th>Med(medreldiff)</th>
<th>Med(1/2IP68)</th>
<th>Pearson R vs CMH; vs CTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLOUD v1; high</td>
<td>21</td>
<td>-2 km</td>
<td>-30%</td>
<td>2.2 km</td>
<td>0.47; 0.28</td>
</tr>
<tr>
<td>CLOUD v1; low</td>
<td>21</td>
<td>-0.3 km</td>
<td>-14%</td>
<td>0.8 km</td>
<td>0.29; 0.32</td>
</tr>
<tr>
<td>CLOUD v2; high</td>
<td>13</td>
<td>-3 km</td>
<td>-38%</td>
<td>1.9 km</td>
<td>0.53; 0.52</td>
</tr>
<tr>
<td>CLOUD v2; low</td>
<td>13</td>
<td>-0.4 km</td>
<td>-14%</td>
<td>0.6 km</td>
<td>0.56; 0.53</td>
</tr>
<tr>
<td>CH: CRB-CLOUDNET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLOUD v1; high</td>
<td>21</td>
<td>-0.9 km</td>
<td>-20%</td>
<td>1.5 km</td>
<td>0.54; 0.35</td>
</tr>
<tr>
<td>CLOUD v1; low</td>
<td>21</td>
<td>-0.4 km</td>
<td>-36%</td>
<td>0.6 km</td>
<td>0.25; 0.21</td>
</tr>
<tr>
<td>CLOUD v2; high</td>
<td>13</td>
<td>-1.2 km</td>
<td>-26%</td>
<td>1.8 km</td>
<td>0.48; 0.51</td>
</tr>
<tr>
<td>CLOUD v2; low</td>
<td>13</td>
<td>-0.6 km</td>
<td>-40%</td>
<td>0.5 km</td>
<td>0.47; 0.51</td>
</tr>
<tr>
<td>CH: FRESCO-CLOUDNET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRESCO v1.3; high</td>
<td>22</td>
<td>-0.6 km</td>
<td>-14%</td>
<td>1.9 km</td>
<td>0.43; 0.32</td>
</tr>
<tr>
<td>FRESCO v1.3; low</td>
<td>22</td>
<td>-0.4 km</td>
<td>-26%</td>
<td>0.6 km</td>
<td>0.35; 0.43</td>
</tr>
<tr>
<td>FRESCO ≥v1.4; high</td>
<td>13</td>
<td>-0.1 km</td>
<td>-3%</td>
<td>1.8 km</td>
<td>0.33; 0.39</td>
</tr>
<tr>
<td>FRESCO ≥v1.4; low</td>
<td>13</td>
<td>0.4 km</td>
<td>29%</td>
<td>0.6 km</td>
<td>0.22; 0.36</td>
</tr>
</tbody>
</table>
L2_CLOUD CRB cloud height is generally below the CLOUDNET cloud mean height. This can be seen in Figure 115, which presents boxplots per station of absolute-scale and relative difference (top panels) and of the monthly median difference (bottom panel). The initially positive bias at Ny-Ålesund is reduced with the upgrade to CLOUD v2. Note that the comparison of CLOUD CRB v1 CH vs CLOUDNET CH at Iquique is impacted by outlying values (as can be seen from the long tail in the boxplot in Figure 115). While there are no CLOUDNET co-locations available for CLOUD CRB v2, we note that the agreement at Iquique of CLOUD CRB with FRESCO CH improved strongly with the introduction of v2.
To obtain an overall value for the bias and the relative bias, we calculate the median over all per-station medians of the difference (CRB CH-CLOUDNET CH) and of the relative difference. Both for version 1 and 2, values of -1 km and 20-25% (rounded to 5%) are obtained for high clouds, and -0.5 km and -40% for low clouds (see Table 17). Note that the number of stations used for version 2 is more limited.

Figure 116 displays boxplots of the difference and relative difference between S5P FRESCO CH and CLOUDNET CH, and the monthly medians per station. The shift towards positive values upon the introduction of FRESCO 1.4 is clear. For version 1.3, bias (relative bias) of FRESCO vs CLOUDNET CH are -0.6 km (-15%) for high clouds, and -0.4 km (-25%) for low clouds. For version 1.4 and later, the bias (relative bias) is now about zero for high clouds, and 0.4 km (30%) for low clouds.

Figure 115: Upper panel: Boxplots of S5P L2_CLOUD CRB RPRO+OFFL 1.1,7-2.3 CH minus CLOUDNET CH (upper left) and of the relative difference (upper right), per station. The same conventions as for Table 17 apply. Sensing time range is indicated on the figure. Bottom panel: monthly median of CLOUD CRB CH-CLOUDNET CH, per station. Pixel size switch (2019/08/06) and UPAS version changes (v1 to v2.1: 2020/07/13, v2.1 to v2.2: 2021/07/01, v2.2 to v2.3: 2022/03/06) are indicated on the plot.
Figure 116: Upper panel: Boxplots of S5P L2 CLOUD FRESCO RPRO+OFFL v1.3-v1.4 CH minus CLOUDNET CH (upper left) and of the relative difference (upper right), per station. The same conventions as for Table 17 apply. Sensing time range is indicated on the figure. Bottom panel: monthly median of FRESCO CH-CLOUDNET CH, per station. Pixel size switch (2019/08/06) and FRESCO version changes (v1.4: 2020/11/29, v2.2: 2021/07/01, v2.3: 2021/11/17) are indicated on the plot.
12.3.4.2 Dispersion

Comparison with alternative S5P cloud height retrievals

The comparison of S5P CLOUD CRB CH vs S5P FRESCO CH reveals a low \( \frac{1}{2} \) IP68 close to the dispersion requirement of 0.5 km at most stations (Figure 117, right). Exceptions are Iquique, Granada, Munich, Yacanto, Ny-Ålesund and Punta-Arenas where the dispersion is higher.

Figure 117 shows the evolution of monthly \( \frac{1}{2} \) IP68 with time. The upgrade of CLOUD CRB to v2.1 causes a decrease of the dispersion at Iquique, Mindelo, Punta-Arenas. Note we have seen earlier that also the median difference CRB-FRESCO at Iquique was reduced with this upgrade. Since the introduction of CLOUD CRB v2.2 and FRESCO v2.2 (both at 2021/07/01) more dispersion reductions seem visible: at Munich, Iquique, and Yacanto.

The dispersion of CAL CTH vs FRESCO CH is typically higher, exceeding the dispersion requirement. This is likely related to the different model assumptions (layer model vs Lambertian model). Still, it is interesting to note that dispersions exceeding 1 km are strongly reduced upon the introduction of CLOUD CAL v2.1.

Figure 117: Top. Left: Monthly \( \frac{1}{2} \) IP68 of the cloud height difference of CLOUD CRB OFFL vs FRESCO OFFL, per month and per station. Indicated are the pixel size switch (2019/08/06), the CLOUD CRB v1 to v2 upgrade (2020/07/13) the FRESCO v1.3 to 1.4 upgrade (2020/11/29), the upgrade from CLOUD CRB v2.1 to v2.2 and from FRESCO v1.4 to v2.2 (both on 2021/07/01), the FRESCO v2.2 to 2.3 upgrade (2021/11/14) and the upgrade from CLOUD v2.2 to v2.3 (2022/03/06). Right: \( \frac{1}{2} \) IP68 of S5P CLOUD CRB CH minus S5P FRESCO CH. Sensing date range is indicated on the figure. Bottom. Similar but for CLOUD CAL CTH vs FRESCO CH.
Comparison with CLOUDNET cloud top height and cloud height

From the width of the boxplots in Figure 114 to Figure 116, it can be inferred that the dispersion of S5P CLOUD CAL CTH minus CLOUDNET CTH, S5P CLOUD CRB CH minus CLOUDNET CH and S5P FRESCO CH minus CLOUDNET CH exceeds the upper limit for error dispersion (500 m). However, also CLOUDNET CTH random error, and comparison error, contribute to the difference dispersion, and these contributions have not been quantified yet.

Figure 118 displays the monthly dispersion per station of S5P CLOUD CAL CTH minus CLOUDNET CTH, S5P CLOUD CRB CH minus CLOUDNET CH, and FRESCO CH minus CLOUDNET CH. No strong evolution is visible, except the decrease of dispersion at Ny-Ålesund and at Punta-Arenas for S5P CLOUD CAL CTH minus CLOUDNET CTH with the upgrade of CLOUD from v1 to v2.

![Figure 118. Monthly dispersion (½IP68) per station of S5P CLOUD CAL CTH minus CLOUDNET CTH (top), S5P CLOUD CRB CH minus CLOUDNET CH (bottom left) and S5P FRESCO CH minus CLOUDNET CH (bottom right). Indicated are the pixel size switch (2019/08/06), the CLOUD v1 to v2 upgrade (2020/07/13), the FRESCO v1.3 to 1.4 upgrade (2020/11/29) and the upgrade from CLOUD v2.2 to v2.3 (2022/03/06).](image)

Overall values of dispersion are obtained by taking the median over all per-station ½ IP68 values (see Table 17). Distinguishing between high (CLOUDNET CTH >4km) and low (CLOUDNET CTH<4 km) clouds, and between version 1 and version 2,

- the dispersion of CLOUD CAL CTH vs CLOUDNET CTH is 2 km for high clouds (version 1 and version 2), and 0.8 km for low clouds for version 1, and 0.6 km for version 2.
- For CLOUD CRB CH, the difference dispersion with CLOUDNET CH is ~1.5 km for high clouds and ~0.5 km for low clouds (both v1 and v2).
• For FRESCO CH the dispersion is ~1.8 km for high clouds for v1.3 and ≥v1.4. For low clouds the dispersion is ~0.6 km.

The Pearson-R correlation over all the individual measurement pairs was calculated, testing both CLOUDNET CH and CLOUDNET CTH as the independent variable, and making distinction between v1 and v2 (for CLOUD OCRA-ROCINN) and v1.3 and ≥v1.4 (for FRESCO) and between low and high clouds (see Table 17). Conclusions are as follows:
• The correlation of CLOUD CAL CTH vs CLOUDNET CH and CTH improves with the introduction of v2; very markedly for low clouds (from ~0.3 to ~0.5).
• The correlation of CLOUD CRB CH vs CLOUDNET CH and CTH improves in most cases with the introduction of v2; very markedly for low clouds (from ~0.2 to ~0.5).
• The correlation of FRESCO CH vs CLOUDNET CH and CTH decreases in most cases with the introduction of v1.4. The exception is for high clouds, when using CLOUDNET CTH as the independent variable. Correlation coefficients vary between 0.2 and 0.4.

12.3.4.3 Dependence on influence quantities

Comparison with CLOUDNET cloud top height and cloud height

Above, we have shown that at the station Iquique (a coastal station), higher discrepancies are encountered for CLOUD CRB CH – FRESCO CH (Figure 109, top), and CLOUD CRB CH – CLOUDNET (Figure 115). This can be attributed to a cluster of data points with low cloud fraction, where CLOUD CRB predicts a high CH, while FRESCO, CLOUD CAL and CLOUDNET predict a low cloud (top) height. The issue is resolved with the introduction of CLOUD v2.

12.3.4.4 Short term variability

Nothing to report.

12.3.4.5 Geographical patterns

Cloud parameters in UPAS 1.1.x may show an enhancement at the east edge of the swath for some months at certain latitudes (Compernolle et al., 2021). The effect seems to be strongest in the latitude bands [40-60]N and [30-40]S. This issue is reduced in UPAS 2. An example for CLOUD CAL cloud top height is shown in Fig. 13 of Compernolle et al. (2021).

12.4 Comparison of L2_CLOUD NRTI and OFFL products

This section investigates if L2_CLOUD NRTI and OFFL are significantly different. Before the upgrade to version 02.01.03, CLOUD NRTI and OFFL use the same algorithm and therefore their difference is expected to be negligible. After the upgrade, CLOUD OFFL also incorporates VIIRS cloud mask information while CLOUD NRTI does not, so a difference can be expected.

Comparisons of CLOUD CAL vs CLOUDNET and CLOUD CRB vs CLOUDNET, available at the VDAF-AVS, were intercompared for OFFL and NRTI. There are hardly visible differences (inspection at 2021/12/02; data not shown).
13 Validation Results: L2_AER_AI

13.1 L2_AER_AI products and requirements

This section reports on the validation of the following geophysical variables of the S5P TROPOMI L2_AER_AI UV aerosol absorbing index products identified in Table 1. Validation results are discussed with respect to the product quality targets outlined in Table 3. The NRTI and OFFL processors producing very similar data products, only validation of the L2_AER_AI NRTI product is reported hereafter. Subsection 13.4 demonstrates evidence that NRTI and OFFL data do not differ significantly and that their respective validations yield similar conclusions. The bias caused by the degradation of the level 1 irradiance data was corrected after processor upgrade in July 2021.

13.2 Validation approach

The UV aerosol index (UVAI) is not a geophysical quantity that can be directly compared to independent measurements from ground or to model results. The way to validate this index is to compare it to coincident satellite measurements from different sensors. For the validation of S5P TROPOMI UVAI, measurements from EOS-Aura OMI and Suomi-NPP OMPS are well suited for that purpose.

In addition to the validation using satellite observations, the S5P TROPOMI UVAI data products can also be checked for internal consistency. For example, the following tests can be performed:

a) the dependence of the UVAI on the observation geometry (in particular on the SZA and the VZA of the measurement) can be investigated;
b) the UVAI values for clear sky and low aerosol amount should be close to zero;
c) the geographical patterns of the UVAI can be compared to those of other measurements, e.g., trace gas distributions of large biomass burning plumes or volcanic plumes.

It should be noted that for S5P TROPOMI the UVAI is calculated for two wavelength pairs, 388 / 354 nm and 380 / 340 nm, the first one allowing a direct comparison to the UVAI from OMI (which is also calculated for 388 / 354 nm).

13.2.1 Ground-based networks

As stated above, satellite UVAI data cannot be directly compared to ground-based measurements.

13.2.2 Satellites

S5P TROPOMI UV aerosol index data are compared to the aerosol indices obtained from EOS-Aura OMI and Suomi-NPP OMPS. Both OMI and OMPS have similar afternoon overpass times as compared to TROPOMI. With OMI the same wavelength pair (388 / 354 nm) can be compared.

13.2.3 Field campaigns and modelling support

As stated above, no direct comparison of the UVAI to non-satellite measurements is possible.
13.3 Validation of L2_AER_AI NRTI

13.3.1 Recommendations for data usage followed

In order to avoid misinterpretation of the data quality and to avoid the effects of sun glint, it is recommended to only use those TROPOMI pixels associated with a qa_value above 0.8. The variables aerosol_index_340_380_precision and aerosol_index_354_388_precision can also be used to diagnose the quality of the UVAI. These are new data product fields and are under evaluation.

For further details, data users are encouraged to read the Product Readme File (PRF), Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) associated with this data product, all available on https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms[ER_CoperATBD].

13.3.2 Status of validation

This section presents updated validation results obtained as a part of the S5P Mission Performance Centre (MPC) and by S5P Validation Team (S5PVT) AO projects. It is based on regular updates of the results reported at the S5P First Public Release Validation Workshop (ESA/ESRIN; June 25-26, 2018) and at the 3rd SPPVT workshop (ESA/ESRIN; November 11-14, 2019). Individual contributions to the workshop are archived in https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5p-first-product-release-workshop/sentinel-5p and in https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5p-precursor-workshop-2019/sentinel-5p, respectively.

The validation of S5P TROPOMI L2_AER_AI data presented here is based on comparisons with similar aerosol indices from the EOS-Aura OMI and Suomi-NPP OMPS satellite missions. Both OMI and OMPS have similar afternoon overpass times as compared to TROPOMI and with OMI the same wavelength pair (354/388 nm) can be compared. Focus is placed on several case studies for different known aerosol sources using reprocessed data from the period covered during the E1 Commissioning Phase (November 2017 to April 2018). The typical case studies identified in Table 18 were selected to cover different types of aerosol plumes expected to be detected by TROPOMI: biomass burning smoke, desert dust, and volcanic aerosol sources. One example for desert dust is shown in Figure 119.

Table 18 – Case studies for different aerosol types.

<table>
<thead>
<tr>
<th>Date</th>
<th>Type of case</th>
<th>TROPOMI orbit</th>
<th>OMI orbit</th>
<th>OMPS orbit</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017-11-10</td>
<td>Desert dust and small Sub-Saharan fire plumes</td>
<td>00398</td>
<td>70864</td>
<td>31285</td>
</tr>
<tr>
<td>2017-11-27</td>
<td>Volcanic eruption, Bali</td>
<td>00636</td>
<td>71108</td>
<td>31523</td>
</tr>
<tr>
<td>2017-12-13</td>
<td>Large biomass burning fires, California</td>
<td>00858</td>
<td>71350</td>
<td>31745</td>
</tr>
<tr>
<td>2018-03-31</td>
<td>Long-range transport of large desert dust plumes</td>
<td>2397, 2398</td>
<td>72916, 72917</td>
<td>33284, 33285</td>
</tr>
</tbody>
</table>
Figure 119: Comparison of S5P TROPOMI UVAI (orbit 00398, left) and OMI OMAERO UV Aerosol Index (orbit 70864, right) for Saharan dust on 10 November 2017. In general very good agreement is found (the stripes in north-south direction in the OMI data are caused by the OMI row anomaly and should be ignored).

For the selected case studies, in general very good agreement of the patterns of enhanced UVAI was found. Comparison results between S5P TROPOMI and OMPS UVAI are shown in Figure 120 and Figure 121 below (courtesy of Omar Torres and Changwoo Ahn, NASA-GSFC). At the beginning of TROPOMI measurements (Nov. and Dec. 2017, Figure 120), the patterns of enhanced UVAI agree very well. But the S5P TROPOMI UVAI is mostly negative and is systematically smaller than the OMPS results. From the start of the operational data record (30 April 2018), a steadily increasing negative bias of the S5P TROPOMI UVAI was present where starting in March 2019; the UVAI was outside the requirements (bias < 1 UVAI unit). The spread of the S5P TROPOMI values is similar as the OMPS values (assuming LER clouds). From this finding, it is concluded that the S5P TROPOMI UVAI is also within the requirement for random errors of 0.1 UVAI units. It should be noted that the standard deviation of the OMPS Mie product is systematically smaller due to the more realistic assumptions about clouds and surface reflectance. A second comparison is performed for measurements in August 2018 (Figure 121). Here again, the spatial patterns agree very well. However, the S5P TROPOMI observations now show systematically decreased UVAI values, which are mostly outside the requirements (bias < 1 UVAI unit). This may in part be related to a wavelength dependent degradation in the irradiance measurements where, shorter wavelengths are more affected. In addition, the spread of the S5P TROPOMI UVAI values has become broader than during the early phase of measurements (see Figure 124). In addition, the reason for this degradation of the data quality has to be further investigated.
Desert dust, 10 Nov 2017

Biomass burning, 12 Dec 2017

Figure 120: Comparison of UVAI from TROPOMI and OMPS for a situation with desert dust (10 Nov 2017, top) and biomass burning (12 Dec 2017, bottom). For OMPS, UVAI are calculated either assuming LER or Mie clouds. The UVAI for Mie clouds yield results that are more consistent. The frequency distributions indicate that S5P TROPOMI results have a similar distribution as the OMPS UVAI calculated for the LER assumption. But TROPOMI values are systematically smaller than the OMPS values (courtesy of Omar Torres and Changwoo Ahn, NASA-GSFC).
Figure 121: Comparison of UVAI from TROPOMI and OMPs for an observation of a biomass burning plume (18 Aug. 2018). For OMPS UVAI are calculated assuming either LER or Mie clouds. The UVAI for Mie clouds yields results that are more consistent. In contrast to early TROPOMI observations, values have systematically decreased and the spread of the UVAI values has become larger (courtesy of Omar Torres and Changwoo Ahn, NASA-GSFC).
In addition, comparisons with patterns from other S5P TROPOMI products are performed. Figure 122 below shows an example of measurements of UVAI and NO₂ VCDs, for which enhanced NO₂ and UVAI are found at the same locations.

Figure 122: Comparison of NO₂ VCDs (left) and UVAI (right) obtained from S5P TROPOMI for sub-Saharan fires on 10 November 2017.

From the performed validation studies it is concluded that the L2_AER_AI UVAI from S5P TROPOMI is of very good quality and fulfilled the requirements until early 2019. The negative bias found in the S5P TROPOMI data, which continues to increase systematically is outside the bias requirements (+/- 1 UVAI unit) since the beginning of 2019. Here it should be noted that the bias is caused by the degradation of the level 1 irradiance data. It was corrected after processor upgrade in July 2021. Also, the spread of the UVAI should be further investigated. Investigations are underway to possibly improve this spread by using a more realistic cloud model (Mie) and surface reflectance.

13.3.3 Bias
The systematic difference between S5P TROPOMI and other instruments measuring aerosol index (OMI and OMPS) was within the requirements earlier in the mission: bias < 1 UVAI unit. Comparisons based on the case studies listed in Table 18 above conclude to a mean bias of -0.8990 AAI with OMPS (TROPOMI UVAI 354/388 – OMPS LER AI 340/378.5). Starting in 2019, the UVAI was slightly outside (below 1 UVAI unit). Currently, with version 2.2.0 data, the UVAI is once again compliant to bias requirements with a global mean average close to -0.5 UVAI units.

13.3.4 Dispersion
The S5P TROPOMI UVAI is very probably within the requirement for random errors of 0.1 UVAI unit. But this preliminary conclusion needs further investigation and confirmation.

13.3.5 Dependence on influence quantities
There is a slight cross-track dependence of -0.25 (West – East side of TROPOMI swath), which is related to the use of the LER model in the retrieval. It should be noted that this cross-track dependence decreases with increasing UVAI values. This finding needs further investigation too.

13.3.6 Short term variability
The global mean aerosol index is evaluated to give an overall indication of the stability of the data product. The global mean is calculated for all pixels on day with full global coverage and it is not expected to vary from day-to-day. A time series of the global mean is given for the TROPOMI UVAI for both wavelength pairs and for the NRTI and OFFL data streams. The period of 20 July 2018 through October 2021 is shown in Figure 123, as the NRTI data coverage was only adequately complete starting 20 July 2018.
The global mean is more negative for the 340/380 wavelength pair as compared to the 354/388 pair. In general, the values for both pairs are more negative than OMI and OMPS global mean averages. This may in part be related to a wavelength dependent degradation in the irradiance measurements where, shorter wavelengths are more affected. This is also most likely why the 340/380 pair is more negative than the 354/388 nm pair. The values of the global mean for all four plots show an overall decrease consistent with the overall degradation trend monitored by the L1b team. This degradation in the irradiance is a known feature in the L1b data and has been corrected with the 2.2.0 version of the data. Additionally, an offset has been applied to resultant data that were positively biased so that the global mean of both wavelength pairs is close to the expected AI value of -0.5.

The values of the global mean and median are nearly identical between the NRTI and OFFL data. The differences are typically in the range of 0.01 - 0.1 and fall well within the expected errors of the UVAI. The structure of variability is slightly different but the overall shape is quite similar, where small structure differences are due to differences in global coverage and/or sampling between the two data streams. The structure and variability when comparing wavelength pairs for the same data stream (i.e. 340/380 NRTI vs. 354/388 NRTI) is also nearly identical. From this comparison, it can be drawn that NRTI and OFFL data streams are comparable with only minor differences and that the wavelength pairs vary in a similar way with an absolute difference no larger than 0.3 UVAI units.
Figure 123: Comparison of the global daily mean and median for both L2_AER_AI UVAI wavelength pairs (340/380 and 354/388 nm) and for the NRTI (top) and OFFL (bottom) data streams, from 20 July 2018 through April 2022.
Figure 124: Comparison of the frequency distribution of the UVAI (left: 354/388nm, right: 340/380nm) for five selected days (20 July 2018, 04 February 2020, 23 May 2020, 21 July 2021, and 21 November 2021).
13.3.7 Geographical patterns

There are no obvious geographical features. For pixels (partially) covered by clouds with a small horizontal extent and a non-homogeneous vertical structure, these clouds are non-Lambertian and result in positive values similar to that of absorbing aerosol. It should also be noted that for many fully clouded scenes, aerosols might be located below the clouds and are therefore invisible for the satellite instrument.

13.3.8 Other features

As mentioned above, the (increasing) negative bias and spread of the S5P TROPOMI results should be reduced in further updates. The negative bias has been partially corrected once with the implementation of version 2 of L1b data. Until the reprocessed data is made available, users should take caution when performing any kind of trend analysis. In recent months an asymmetry of the frequency distribution was found, which needs further investigations.

13.4 Equivalence of L2_AER_AI NRTI and OFFL products

**Figure 125** below shows a comparison for a selected orbit on October 3, 2018. For this orbit, the L2_AER_AI UV aerosol absorbing index for both wavelength pairs are very similar for the OFFL and NRTI products. Based on this comparison and the comparison of the global means shown before, the close similarity in behaviour of both the NRTI and OFFL data streams indicates that the validation results for the NRTI data product are also valid for the OFFL data product.

![Comparison of S5P TROPOMI UVAI for a selected orbit (#05033) on 3 October 2018 for the two wavelength pairs (top: 340 / 380 nm, bottom: 354 / 388 nm). While the geographical patterns are the same, the absolute values differ slightly with the NRT values (left) slightly higher than the offline values (right).](image-url)
14 Validation Results: L2_AER_LH

14.1 L2_AER_LH products and requirements

This section reports on the validation of the S5P TROPOMI L2_AER_LH aerosol layer height (ALH) product as identified in Table 1. Validation results are discussed with respect to the product quality targets outlined in Table 3. Only validation of the L2_AER_LH OFFL product is reported here.

14.2 Validation approach

The Aerosol Layer Height was released to the public on September 30, 2019 and is still a relatively new product. The validation of the product is ongoing. The results presented here represent the reprocessed 2018 data (OFFL) that were generated after an extensive update of the product, in which the forward component of the algorithm fit was updated with a neural network predicting the reflectance and derivatives of the radiative transfer model. This allows global processing of data in near-real time, while still using the same optimal estimation approach. The reprocessed data have been validated before release, which is what is presented here.

The ALH presented here is computed only for known aerosol layer, which, lacking an AOT product, is done by selecting high UV AI values (larger than 0). This means that mainly desert dust, smoke and volcanic plumes will be processed. Therefore, the validation focused on selected desert dust cases, fires plumes and occasional volcanic eruptions. In upcoming releases of the ALH product, the ALH will be computed for all cloud-free pixels, regardless of the UV AI value.

Furthermore, since no global aerosol layer height products are available next to TROPOMI’s ALH, the validation is limited to co-locations with satellite observations: the MISR’s stereoscopic layer height product, and CALIOP’s active sensing of the atmospheric vertical profile. Both instruments have a limited swath, therefore finding suitable co-location is the main limiting factor for intercomparison.

14.2.1 Ground-based networks

Validation of the TROPOMI ALH with ground-based networks is desirable, since satellite-to-satellite comparisons have their own specific limitation, as stated above. Validation of the TROPOMI ALH with ground-based networks was carried out using several EARLINET stations in the Mediterranean region.

The European Aerosol Research Lidar Network, EARLINET (https://www.earlinet.org), was founded in 2000 as a research project for establishing a quantitative, comprehensive, and statistically significant database for the horizontal, vertical, and temporal distribution of aerosols on a continental scale (Pappalardo et al., 2014). Since then EARLINET has continued to provide the most extensive collection of ground-based data for the aerosol vertical distribution over Europe. EARLINET observations are performed on a regular schedule for daytime and nighttime measurements. In addition to these systematic measurements for the consolidation of a European aerosol climatology, further observations are devoted to monitoring special events over the continent, such as Saharan dust outbreaks, forest fires, photochemical smog, and volcanic eruptions. The EARLINET database represents the largest collection of ground-based data of the vertical aerosol distribution on a continental scale. The main information stored in the files of the EARLINET database is the vertical distribution backscatter and aerosol extinction coefficients. The basic issue in this validation approach is the difficulty in identifying good spatiotemporal collocations between EARLINET lidar stations observations and TROPOMI/S5P overpasses.
14.2.2 Satellites

S5P TROPOMI aerosol layer height data were compared to the stereoscopic plume height product from MISR and to the weighted extinction height provided by CALIOP. The stereoscopic plume height product from MISR is an offline product that can be computed for selected fire plumes, using a freely available code (MINX). It makes use of the nine available viewing directions of MISR, which senses a scene from different directions during an overpass. This provides stereoscopic height information for a scene with enough contrast. The MINX code has to be processed manually, and also the fire plumes have to be hand-picked and selected digitally by hand. In this document plumes from 115 fires in 2018, prepared and provided by D. Griffin from the Environment and Climate Change Canada institute, are compared with TROPOMI ALH. Furthermore, the weighted extinction height from CALIOP on Calipso are compared to TROPOMI ALH for collocated pixels. All pixels were selected where Calipso was closer to S5P than 100 km and the sensing time of CALIOP and TROPOMI was less than three hours apart. The resulting number of pixels (about 2.5 million pixels in from May 2018 – March 2019) were screened for clouds and selected for aerosols. This resulted in about 1 million pixels over the oceans and 0.5 million pixels over land. The results of the comparisons are presented below.

A few more satellite products are available for comparison with the TROPOMI ALH. GOME-2 provides the Absorbing Aerosol Height (AAH), which is layer height product that is computed for selected pixels with high UV-AI, representing thick absorbing aerosol plumes. The AAH is comparable to the ALH since it also uses the depth of oxygen absorption lines in the O2-A band to derive the height of scattering layer. However, it differs from the ALH in that it only uses one or a few absorption lines and the continuum, while the TROPOMI ALH fits about 3,500 lines in the O2-A band, which should make it more accurate than the AAH. A similar product as the GOME AAH is available from EPIC on DSCVR. This product can be expected to have similar accuracy as the GOME AAH, but since DSCVR is parked in Lagrangian point L1 between the Sun and the Earth, it can deliver aerosol layer height at a one-hour time resolution. This would make it possible to monitor the evolution of aerosol layer heights, and cover the time differences between overpasses of e.g. Calipso and MIRS, and TROPOMI.

14.2.3 Field campaigns and modelling support

So far, no field campaigns have been planned to validate the ALH.

14.3 Validation of L2_AER_LH

14.3.1 Recommendations for data usage followed

The ALH is very sensitive to cloud contamination. However, aerosols and clouds can be difficult to distinguish, and ALH is computed for all FRESCO effective cloud fractions smaller than 0.05. Since the ALH is sensitive to elevated scattering layers, and cloud layers are generally optically (much) thicker than aerosol layers, not discriminating between clouds and aerosol will strongly bias the ALH towards cloud layer heights. Cloud masks are available from FRESCO and VIIRS, and are strongly recommended to filter for residual clouds. A sun-glint mask is also available to screen sun-glint regions, which are not filtered beforehand. These and other sources of uncertainties are indicated with the qa_value. Use of pixels with a qa_value below 0.5 is not recommended.

The variables aerosol_mid_pressure_precision and aerosol_mid_height_precision can also be further used to diagnose the quality of the ALH.
For further details, data users are encouraged to read the Product Readme File (PRF), Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD) associated with this data product, all available on https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms [ER_CoperATBD].

14.3.2 Status of validation

This section presents validation results obtained as a part of Validation Team (S5PVT) AO projects and development tests during the update of the forward model.

The validation of S5P TROPOMI L2_AER_LH data presented here is based on comparisons with ground based observations (see section 13.2.1) as well as MISR and CALIOP, as detailed in 14.2.2. In Table 19, the details of four selected cases for the satellite to satellite comparisons are presented, which were compared to the CALIOP weighted extinction height. A fifth case of very high altitude smoke from intense biomass burning in Australia in early 2020 shows a notable difference with CALIOP measurements, showing a limitation of the S5P L2_ALH product. The comparison results of S5P TROPOMI L2_AER_LH to ground based observations are presented at the end of this section.

Table 19 – Case studies for desert dust cases with Calipso co-locations

<table>
<thead>
<tr>
<th>Date</th>
<th>Type of Case</th>
<th>TROPOMI orbit</th>
<th>Calipso orbit start time [Day/Night]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018-06-01</td>
<td>Desert Dust</td>
<td>3280</td>
<td>14:28:34 [D]</td>
</tr>
<tr>
<td>2018-06-08</td>
<td></td>
<td>3379</td>
<td>14:34:52 [D]</td>
</tr>
<tr>
<td>2018-06-10</td>
<td></td>
<td>3407</td>
<td>14:22:32 [D]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3408</td>
<td></td>
</tr>
<tr>
<td>2018-12-22</td>
<td>Smoke</td>
<td>7163</td>
<td>12:55:29 [D]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7174</td>
<td></td>
</tr>
<tr>
<td>2020-01-11</td>
<td></td>
<td>11640</td>
<td>07:54:18 [N]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11641</td>
<td></td>
</tr>
</tbody>
</table>
Figure 126: Details of the selected validation cases, showing the ALH on a VIIRS RGB background. The black line represents the Calipso track.

Figure 126 show the cases, (a)-(c) are similar desert dust cases, with dust blowing off the African continent over the Atlantic, and (d) is a smoke case, with smoke over both land and ocean. The black lines display the CALIOP tracks, which have a good coverage of the events. The curtain plots from CALIOP are shown in Figure 127, displaying the total attenuated backscatter as measured by CALIOP in the colour code shown next to the plot, the weighted extinction height in black-and-white, and the ALH from collocated TROPOMI pixels in blue-and-white.

First, the images show that the maximum attenuated backscatter measured by CALIOP, at 532 nm, is not a good indicator of the plume centre height. The maximum total attenuated backscatter is often at the plume top. The weighted extinction height is computed from the level-2 aerosol extinction profiles. Here, aerosol extinction is computed in cloud-free areas using a feature mask, distinguishing (among others) aerosol and cloud layers. Each well-defined aerosol layer and aerosol-free layer is split in 100 m height segments to allow for averaging over complex layer structures along the CALIOP path. The average extinction height is then computed by (Nanda, et al., 2018):
where $Z_i$ is the height from sea level in the $i$th lidar vertical level (in km), and $b_{ext,i}$ is the aerosol extinction coefficient (in km$^{-1}$) at the same level.

The weighted extinction height is an indication of the maximum of the extinction, and is often related to the centre of a plume if the attenuation of the beam is small. However, for strongly attenuated beams, the weighted extinction height is biased to the top of the plume. Figure 127 shows that the weighted extinction height correlates rather well with the TROPOMI ALH. However, TROPOMI generally shows lower altitude plumes heights than CALIOP. Also, clouds strongly bias the TROPOMI ALH towards the cloud altitude. Therefore, additional cloud screening, which is available in the product in the form of flags, is essential for the user to retrieve proper aerosol layer heights.

Figure 127: Curtain plots from CALIOP, showing the total attenuated backscatter at 532 nm for the four cases in Figure 126. Plotted on top of the coloured background image of the total attenuated backscatter are the weighted extinction heights as derived from the backscatter coefficient in black-and-white, and the ALH from collocated TROPOMI pixels in blue-and-white.
In Figure 128, the cases are compared pixel to pixel. Obviously, the four cases represent desert dust and smoke plumes, which may be more or less homogeneously distributed in the atmosphere. Therefore, from the individual cases a linear regression is not meaningful. However, since the layers in the four cases are each at different (average) altitudes, they can be used for a linear regression. This shows a very similar sensitivity of CALIOP and TROPOMI ALH (slope is 1.00), but there is clearly a persistent offset between the two parameters. CALIOP weighted extinction is on average about 0.53 km higher in altitude than TROPOMI ALH. This is likely more due to the differences in method and measured quantities than to systematic errors in the data products themselves.

![Figure 128](image.png)

**Figure 128:** Comparison of ALH from TROPOMI and CALIOP for the cases presented in Figure 126. Each case is colour coded. CALIOP weighted extinction height is consistently lower than TROPOMI ALH.

The comparison between CALIOP and TROPOMI was extended to all collocated pixels within 100 km and within 3 hours of each other, yielding about 1 million pixels over the ocean and 0.5 million over land, see Figure 129 (left). The figure shows that the TROPOMI ALH is systematically lower than CALIOP weighted extinction heights. The retrieved ALH from TROPOMI differs from CALIOP weighted extinction height by 1.0 km on average, with a standard deviation of 1.97 km. More than 50% of the TROPOMI ALH retrievals over the ocean have an absolute difference with CALIOP weighed extinction height by less than 1.0 km. Retrievals over land have a larger difference, with -2.41 km on average and a median of -1.75 km. The results are very skewed over land, with very large negative values dictating the average — this is indicated by the very large standard deviation of 3.56 km. 50% of the selected colocations over land have an absolute difference with CALIOP weighted extinction height less than approximately 1.0 km. On the right, a similar histogram is shown, but now for only those pixels that have a minimal cost function, or χ², smaller than 1E5. The χ² represents the goodness-of-fit of the modelled sun-normalised radiances to the observations in the O2-A band, and therefore is a measure of the representativeness of the model (of a simple one aerosol layer atmosphere with known surface reflectance) to reality. Smaller χ² indicate a better fit. The retrievals over land generally have much higher χ², and therefore are less reliable. The right panel in Figure 129 show the results for pixels with a χ² less than 1E5, which can be expected to be a reasonably good fit. The differences between TROPOMI ALH and CALIOP weighted extinction height then reduce to -0.62 km over ocean and -1.2 km over land.
Figure 129: Histogram of differences between CALIOP weighted extinction height and TROPOMI ALH from collocated data between 1 May 2018 and 28 February 2019 (left). The right panel shows the same histogram, but for pixels that were screened for a minimal cost function (chi-squared) smaller than 1E5.

Additional validation of the TROPOMI ALH was provided by Environment and Climate Change Canada. TROPOMI ALH was compared to MISR stereoscopic plume height and CALIOP “layer_base_altitude” and “layer_top_altitude” products for 115 fire plumes in 2018 over northern America (Griffin et al, 2019). The results are summarized in Figure 130 and Figure 131. The maximum plume heights above ground level for the 2018 fires in North America are, on average, 2 km (ranging between 0.4 and 5.5 km) and 1.6 km (ranging between 0.01 and 8.4 km) for MISR and TROPOMI, respectively. The mean plume heights (above ground level) within one fire plume are on average 1.4 km (ranging between 0.3 and 3.2 km for MISR) and 0.8 km (ranging between 0.01 and 2.8 km for TROPOMI). Overall, TROPOMI’s maximum and mean plume height is on average 0.59±1.3 km and 0.55±0.74 km lower than the plume height derived from MISR, respectively.

The difference between the plume height observed by TROPOMI and CALIOP depends significantly on the thickness of the plume (as derived from CALIOP). Thicker plumes seem to be better captured by TROPOMI and the thicker the plume the smaller the difference between the CALIOP and TROPOMI plume height. TROPOMI was biased low in comparison to CALIOP for thin smoke plumes (thickness of less than 1.5 km) and TROPOMI ALH is on average 2.1 km lower. Much better agreement and a higher correlation between the two satellite datasets is found for thicker plumes. The mean difference reduces with the thickness of the plumes, the mean difference between the TROPOMI and CALIOP mid aerosol layer is just 50 m for very thick plumes (>3 km). The geometrically thick plumes are typically optically thicker plumes, too. The reason for the reduced bias with increasing layer thickness is probably the sensitivity of the TROPOMI AER_LH algorithm to the scattering layer in the scene, which is more and more dominated by the surface if the aerosol layer is optically thinner. Currently, a simple Lambertian Equivalent Reflection (LER) database from GOME-2 is used in the ALH retrieval to fit the observations to the simulated reflectances. An improvement is expected when a (directional) LER database from TROPOMI becomes available.
Figure 130: Comparison of TROPOMI ALH and MISR plume height for 115 fires over Northern America in 2018. See Griffin et al, 2019 for details.

Figure 131: Comparison of TROPOMI ALH and CALIOP average aerosol layer height (top minus bottom of aerosol layer as defined by the feature mask) for collocated pixels near fires over Northern America in 2018. See Griffin et al, 2019, for details.
The validation of S5P TROPOMI L2_AER_LH data presented below is based on comparisons with ground-based lidar stations belonging to EARLINET. EARLINET data from 7 stations were considered in this analysis. We used S5P L2_AER_LH (RPRO+OFFL) 01.03.00-02.03.01 data. This covers the time period from May 2018 till Sep 2021. The geographical distribution of the selected EARLINET stations depicted in Figure 132 indicates the domain of applicability of the validation results. All participating stations (red circles) operate high-performance multi-wavelength lidar systems. The location of the stations across the Mediterranean basin is an ideal test environment for TROPOMI ALH features due to their proximity to the Sahara Desert and Europe, with frequently observed events of mineral dust and smoke particles. The TROPOMI aerosol layer height product can be examined under a complete set of different atmospheric conditions. We further assessed the capabilities of the ALH product over both land and sea. Over land, the TROPOMI ALH product has decreased detection capabilities than over the sea surfaces since, over bright surfaces, the retrieval algorithm becomes increasingly sensitive to errors in the surface albedo features. This validation was performed by the team from the Aristotle University of Thessaloniki (AUTH). For the routine validation of the S5P/TROPOMI aerosol layer height retrievals, the automated validation server in LAP-AUTH deployed within the QA4EO project (Work Package 2191) collects S5P ALH data and correlative measurements to identify suitable co-locations, compares to the co-located data, and produces S5P data quality indicators. The approach followed is mostly based on the previous expertise and methodology that have been developed and applied in EARLINET for the GOME2 cal/val activities (Michailidis et al., 2021a). Detailed information about the validation methodology and current status of the validation results can be found in Michailidis et al. (2021b).

Figure 132: Geographical distribution of EARLINET ground-based stations for which co-locations with S5P L2 AER_LH data were used (period May 2018 – September 2021).

<table>
<thead>
<tr>
<th>Station</th>
<th>Code</th>
<th>Country</th>
<th>Longitude, latitude, elevation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antikythera</td>
<td>AKY</td>
<td>Greece</td>
<td>23.31°E, 35.86°N, 193m</td>
</tr>
<tr>
<td>Athens</td>
<td>ATZ</td>
<td>Greece</td>
<td>23.78°E, 37.96°N, 212m</td>
</tr>
<tr>
<td>Évora</td>
<td>EVO</td>
<td>Portugal</td>
<td>7.91°W, 38.56°N, 293m</td>
</tr>
<tr>
<td>Granada</td>
<td>GRA</td>
<td>Spain</td>
<td>3.60° W, 37.16°N, 680m</td>
</tr>
<tr>
<td>Lecce</td>
<td>SAL</td>
<td>Italy</td>
<td>18.10°E, 40.33°N, 30m</td>
</tr>
<tr>
<td>Limassol1,2</td>
<td>LIM</td>
<td>Cyprus</td>
<td>33.04°E, 34.67°N, 10m</td>
</tr>
<tr>
<td>Potenza</td>
<td>POT</td>
<td>Italy</td>
<td>15.72°E, 40.60°N, 760m</td>
</tr>
</tbody>
</table>

1Cyprus University of Technology (CUT) [before Oct 2020]
2Leibniz Institute for Tropospheric Research and ERATOSTHENES Centre of Excellence [after Oct 2020]
### 14.3.3 Validation approach

The AUTH team used the aerosol layer height retrieved from ground-based lidar systems within EARLINET to validate the TROPOMI ALH product. TROPOMI observations, co-located with the ground-based EARLINET measurements, are found by selecting all filtered and averaged TROPOMI pixels within a radius of 150 km around each station and with a maximal time difference of 4h. Pixels with an associated quality assurance value of less than 0.5 were excluded. This filter does not remove all pixels considered unusable. Some pixels with unphysically low or high ALH information are still present. Some of the known data quality issues are not covered by the quality flag criterion of 0.5 and additional flags should be applied to the data during the validation analysis reported hereafter. Those issues include the insensitivity to very high altitudes aerosol layers and the treatment of remaining high altitude clouds. The strong possibility of remaining clouds in the TROPOMI field-of-view is one of the reasons why an optimal spatial collocation with the lidar measurement is not achieved for every target pixel. Routine validation is done using the automated ALH validation system operated at LAP-AUTH. The TROPOMI pixel selection scheme and flags applied in the presented validation study, were made following the recommendation by the Product User Manual (PUM) and Algorithm Theoretical Basis Document (ATBD).

As input to the validation processing, we use the lidar backscatter coefficient profiles at 1064 nm (or 532 nm), analysed by the Single Calculus Chain (SCC; https://scc.imaa.cnr.it/) algorithm (D’Amico et al., 2016) for quality-assured measurements. The main aim of SCC is to provide any lidar station with a quality controlled data processing chain to retrieve vertical profiles of important aerosol optical parameters in a fully automatic way. The backscatter files contain at least a profile of the aerosol backscatter coefficient ($m^{-1}sr^{-1}$) derived from the elastic backscatter signal. For the layer height retrieval from the EARLINET products, we used the methodology proposed by Mona et al. (2006) to calculate the weighted backscatter height using the Level-2 backscatter profiles 1064 nm (or 532nm).

The parameter ALHbsc is calculated as the backscatter-weighted height according to the expression:

$$ALH_{bsc} = \frac{\int_{z_i}^{z_n} \beta_{aer,i}(z)dz}{\int_{z_i}^{z_n} \beta_{aer,i}(z)dz}$$

where $\beta_{aer,i}$ represents the aerosol backscatter coefficient ($Mm^{-1}sr^{-1}$) primarily at 1064 nm channel at level-i and $Z_i$ is the altitude (km) of level i for the aerosol profile signal. Based on the above equation, the layer height is calculated from backscatter profiles, symbolized as ALHbsc. The ALHbsc represents an effective ALH weighted by the aerosol backscatter signal at each level and is consistent with ALH as defined in the TROPOMI algorithm. In our work analysis, we applied Eq. 1, to all lidar backscatter profiles collocated to TROPOMI measurements.
The results of the comparison between the TROPOMI and EARLINET-derived aerosol heights for 34 identified collocated cases between May 2018 and Sept 2021, are shown in Figure 133 (upper panel) which shows the scatterplot of TROPOMI ALH against EARLINET ALH_{bsc} for all the common cases (N=34) used for the intercomparison. By defining a weighted height from EARLINET aerosol backscatter profile products (ALH_{bsc}), the quantitative validation at pixels over the selected EARLINET stations illustrates that TROPOMI ALH is consistent with ALH_{bsc}, with a high correlation coefficient R=0.89 and mean bias $-1.017 \pm 0.96$ km over ocean pixels and R=0.89 and $-1.459\pm1.57$ km over both land/ocean pixels, respectively. Ocean-only S5P pixel comparisons are shown in Figure 133 (upper left) and both S5P ocean and land pixels are shown (upper right). The yellow solid line is the linear fit line between the datasets. The colour scale indicates the averaged TROPOMI aerosol index values. The S5P TROPOMI ALH is systematically smaller than the EARLINET results. This can be seen also in Figure 133 (bottom panel) which presents histogram plots of absolute differences. The S5P ALH algorithm appears to provide improved comparisons for ocean pixels. Many factors can play a role in this apparent disagreement between TROPOMI retrievals over land and sea. It is known that high surface albedos negatively influence the ALH, biasing the ALH towards the surface.

**Figure 133:** Upper panel: Scatterplots of TROPOMI against EARLINET data. (left) TROPOMI pixels over ocean and (right) over ocean and land. The color of each scatter point indicates the TROPOMI retrieved UVAI (388/354nm) values, and the error bars of each scatter indicate the spatiotemporal variability of the averaged TROPOMI ALH pixels. Bottom panel: Histogram of the differences between TROPOMI and EARLINET datasets, in red over the ocean pixels and in green for both ocean and land pixels. The co-locations cover the period from May 2018 to August 2021.
Validation statistics for collocated pairs are reported in Table 21 for TROPOMI and EARLINET for correlation coefficient R, slope and intercept of a linear regression and mean and median absolute biases.

**Table 21**: Statistics on the comparison of the common subset of L2_AER_LH and co-located EARLINET stations.

<table>
<thead>
<tr>
<th>TROPOMI Pixels</th>
<th>Slope</th>
<th>Y-intercept</th>
<th>Pearson R</th>
<th>Median (km)</th>
<th>Bias±std (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Over ocean</td>
<td>0.59</td>
<td>0.49</td>
<td>0.89</td>
<td>-0.86</td>
<td>~1.02 ± 0.96</td>
</tr>
<tr>
<td>Over ocean &amp; land</td>
<td>0.26</td>
<td>1.26</td>
<td>0.59</td>
<td>-1.09</td>
<td>~1.46 ± 1.57</td>
</tr>
</tbody>
</table>

### 14.3.4 Case studies

To illustrate the evaluation methodology for the TROPOMI Level-2 ALH, selected pairs of collocated and concurrent TROPOMI and EARLINET lidar observations are presented. Two typical case studies identified in the table below were selected to cover different types of aerosol plumes expected to be detected by TROPOMI: desert dust and biomass burning smoke sources. The selected examples for these events are illustrated in Figure 134 and Figure 135 for a case of desert dust (22 June 2021) and biomass burning (26 October 2020) aerosols.

**Table 22**: Case studies for TROPOMI SSP and EARLINET co-locations

<table>
<thead>
<tr>
<th>Date</th>
<th>Type of case</th>
<th>TROPOMI Orbit</th>
<th>EARLINET station</th>
</tr>
</thead>
<tbody>
<tr>
<td>2021-06-22</td>
<td>Desert Dust</td>
<td>19125</td>
<td>Antikythera, Greece</td>
</tr>
<tr>
<td>2020-10-26</td>
<td>Smoke</td>
<td>15735</td>
<td>Potenza, Italy</td>
</tr>
</tbody>
</table>

The first example refers to a Sentinel-5P overpass of the Antikythera, Greece, on 22 June 2021. During this Saharan dust event the PollyXT NOA system was operating in a 24/7 mode in PANGEA observatory. The TROPOMI ALH retrieved pixels are situated in between the surface and 4 km (Figure 134, left). This map shows that a significant dust plume covered Greece, including the island of Antikythera. The aerosol load was located between 2 – 5 km according to lidar measurements and the sky above the site was cloud-free during the TROPOMI overpass. In Figure 134 (right) the closest in time averaged vertical backscatter profile (11:30–13:00 UTC) is illustrated, measured by the PollyXT lidar system. A clear single layer with large aerosol load was observed according to the backscatter vertical profiles. Accordingly, the TROPOMI ALH spatially averaged values and the EARLINET temporally averaged backscatter coefficient profiles are qualitatively compared. TROPOMI detects this layer height approximately around 2.55±0.36 km (and 2.49±0.41 km) over ocean and both ocean/land, respectively, and the calculated ALH_{bsc} from the lidar profile is 3.01 km. The AER_LH retrievals from TROPOMI are shown in the inbox text. A quite satisfactory agreement between the satellite and ground-based lidar systems is shown for this aerosol scene. It is clear that the center of the assumed layer height by the TROPOMI operational algorithm is slightly lower but very close to the actual layer location calculated as ALH_{bsc} height. This case study highlights the fact that, that under homogeneous, relatively cloud-free conditions, the mean ALH value retrieved by TROPOMI is in good agreement with the calculated height from lidar profile.
Figure 134: Left: TROPOMI ALH retrievals (orbit 19125) over Greece during dust event on 22 June 2021. The red star indicates the position of the Antikythera (PANGEA) station. Right: Lidar backscatter profile at 1064 nm. The horizontal dashed-dotted red line represents the calculate profile centre of mass. The AER_LH retrievals from TROPOMI are shown in the inbox text.

On October 26th, 2020, a smoke plume originated from North America (California) fires spread towards the central Mediterranean. Here, we present a case study during this smoke episode, where a significant aerosol load is observed over central and mainly over the south of Italy. The location of the smoke plume is clearly seen in the TROPOMI ALH (Figure 135, left) during the Sentinel-5P overpass between 11:20-12:20 UTC. The maximum altitude in the AER_LH data is about ~8 km. However, the lidar data detect much higher altitudes for the smoke plumes. In Figure 135 (right), the retrieved vertical profile of the observations with MUSA lidar operated in Potenza (CNR-IMAA), is presented. The closest in time backscatter profile is used in order to extract the ALHbsc and compare against TROPOMI ALH retrievals. The AER_LH retrievals from TROPOMI are shown in the inbox text. Averaged backscatter profiles at 1064 nm, for the time period from 10:13 to 11:40 UTC is shown. Two optical thick layers with a thickness of ~2 km were detected. The dashed-dotted red line denote the profile centre of mass according to the equation provided above. The right plot compares the collocated ALHbsc calculated by lidar using the level-2 backscatter at 1064 nm against TROPOMI ALH mean around Potenza site. TROPOMI detects this layer at 5.54±2.2 km using ocean pixels (and 2.58±2.1 km) while the calculated ALHbsc from the lidar profile places it at 7.8 km. Clearly, the AER_LH is placed much lower than the calculated altitude of the lidar profile. The exact reason for the much lower altitude retrieved by the AER_LH algorithm is not clear; however we should note that the AER_LH algorithm was not created to retrieve ALH at such low air pressures. The ALH pixels over Italy show clear outliers, with very low reported heights, along the inland region. All these pixels over land seem to result in an ALH very close to the ground. This case of very high altitude smoke from intense biomass burning in North America in 2020 shows a notable difference with lidar measurements, revealing a limitation of the S5P L2_AER_LH product.
### 14.3.5 Bias

The systematic difference between S5P TROPOMI ALH and MISR aerosol plume height is about 600 m. This is mostly due to differences in the sensitivity of the instruments and the differences in the algorithms. A difference of about 500 m (lower for TROPOMI) is expected from simulations. TROPOMI ALH is sensitive to the centroid aerosol layer height. Furthermore, TROPOMI ALH is more accurate for thicker plumes, when compared to CALIOP aerosol weighted extinction height. For a 3 km thick plume the difference between CALIOP and TROPOMI layer height decrease to only 50 m. The TROPOMI ALH is well within the requirements of 100 hPa for the bias.

From the comparison with ground based observations, a slightly larger difference is found. By defining a backscatter-weighted aerosol height from EARLINET aerosol backscatter profile products (ALHbsc), the quantitative validation at pixels over the selected EARLINET stations illustrates that TROPOMI ALH is consistent with ALHbsc, with a high correlation coefficient R=0.91 (R=0.59) and mean bias about $-1.017 \pm 0.96 \text{ km}$ ($-1.459 \pm 1.57 \text{ km}$) over ocean and ocean/land pixels respectively.

As a final point, it appears that aerosol layer altitudes retrieved from TROPOMI are systematically lower than altitudes from the lidar retrievals. There is a bias, which is related to the use of the LER model in the retrieval. From the start of the operational data record (30 April 2018), a steadily negative bias of the S5P TROPOMI ALH mainly outside the requirements (bias > 1km). This finding needs further investigation to possibly improve this spread by using a more realistic surface reflectance.

### 14.3.6 Dispersion

The S5P TROPOMI ALH dispersion is large due to cloud contamination and surface effects. With rigorous cloud screening, 50 % of the pixels are already within 1 km of the CALIOP weighted extinction height. Accounting for the expected bias, this is within the requirements of 50 hPa. However, this preliminary conclusion needs further investigation and confirmation.
14.3.7 Dependence on influence quantities

The TROPOMI ALH is strongly dependent on subpixel clouds, and cloud filtering remains essential. The user is strongly encouraged to use all available cloud filters. Until the end of June, the ALH was only processed for UV Aerosol Indices larger than zero. However, the UV AI was biased and degrading, which means fewer ALH pixels were processed as time continued. The number of processed ALH pixels since August 2019 is shown in Figure 139. Ignoring a few isolated episodes with a large number of processed pixels, which were caused by periods of extreme wildfires, a gradual reduction in the number of pixels being processed can be observed until the end of August 2021. This is attributed to the reduction in UV-AI values. With the release of version 2 data in July 2021, the UVAI filter was replaced by a cloud mask filter, which increased the number of processed measurements by more than an order of magnitude. This change includes also scattering aerosols in the ALH processing, instead of only light-absorbing aerosols, The algorithm was not changed, so the validation results remain valid for version 2. However, the use of a cloud mask changed the behaviours of the NRTI and OFFL ALH, since NPP VIIRS cloud mask (VCM) is used for OFFL, but this is not available in NRTI. Therefore, in NRTI FRESCO cloud mask is used, which is not a geometric cloud mask like CMA, which results in differences in the number of cloud-free pixels. This is illustrated in Figure 136 to 138.

The difference between the NRTI and OFFL retrieval was investigated for an arbitrary day, 30 Aug 2021. Figure 136 shows the normal results. It is obvious that the cloud masks behave differently, resulting for this day in 659101 NRTI ALH pixels and 867394 OFFL ALH pixels. The difference is illustrated in Figure 137: The NRTI cloud filter FRESCO filters much more clouds over ocean, which are identified as scattering aerosols by the OFFL cloud filter CMA. In contrast, over land more cloud pixels are filtered by CMA in OFFL, which are not recognized by FRESCO in the NRTI processing. However, all these pixels over land seem to result in an ALH very close to the ground.

In Figure 138 the histograms of the retrieved pixels are compared. Obviously, the histograms are very different, especially for low altitude ALH: The OFFL ALH shows a peak around 2 km, while the OFFL ALH has many ALH retrievals below 2 km, increasing towards the surface. Whether these are low altitude residual clouds or boundary layer aerosols is not clear. Currently, the CMA is considered superior to the FRESCO cloud mask. Figure 138 also shows the OFFL vs the NRTI ALH for all pixels that are not filtered by either cloud filter. Clearly, the differences are small for the majority of pixels and the NRTI, confirming that the OFFL and NRTO ALH processors produce similar results.
Figure 136: S5P/TROPOMI ALH retrieval for 30 Aug. 2021 using the OFFL processor (top), resulting in 867394 processed pixels, and the NRTI processor (bottom), resulting in 659101 processed pixels.
Bright surfaces have a strong effect on the ALH, and very high ALH (altitude up to 12 km) often occur over the Saharan desert. These should be filtered, but a filtering scheme is currently not available. Sun-glint produces high UV-AI values and are processed for ALH. These ALH values show up in overview plot, but are easily filtered using the sun-glint filters. Also, the ALH for aerosol-free sun-glint areas are close to zero (altitude) as expected.

**Figure 137**: Difference between the S5P/TROPOMI NRTI and OFFL ALH retrievals for 30 Aug. 2021 shown in Figure 136.

**Figure 138**: (Left) Histograms of the S5P/TROPOMI NRTI and OFFL ALH retrievals for 30 Aug. 2021 shown in Figure 136; (right) Scatterplot of the OFFL and NRTI ALH retrievals for all pixels that were selected by both the NRTI and OFFL cloud masks.
Figure 139: Number of successfully processed ALH pixels per day from April 2019 to April 2022. In July 2021, a cloud mask filter replaced the UVAI filter.

The comparison to ground based observations indicate that the S5P L2 AER_LH reports unphysically low values over land due to strong illumination. The high surface albedo above land surfaces is a challenge for ALH retrievals. The TROPOMI ALH is expected to decrease with the implementation of version 2. Until the reprocessed data is made available, users should be take caution when performing any kind of analysis.

14.3.8 Short term variability

The high correlation coefficients between TROPOMI and ground based observations demonstrate the ability of the TROPOMI observations to properly capture the temporal variability of tropospheric aerosol plumes. But the TROPOMI ALH data product is strongly event driven and more detailed remarks on the short-term variability cannot be provided at the moment.

14.3.9 Geographical patterns

There are no obvious geographical features.
14.3.10 Other features

A limitation of the S5P ALH product has become apparent following the severe bushfires in New South Wales during the 2019-2020 fire season. Hundreds of severe wildfires have consumed an estimated 18.6 million hectares in the southeast of Australia. The smoke and gases from these fires were well visible in several S5P products, including UV-AI, ALH and HCHO and CO total column. In Figure 140 a screen shot shows the S5P ALH on 11 January 2020 over the south Pacific as displayed on the S5P-TROPOMI-KNMI-Level 2 Product Maps webpage. It shows the extent of the fire ash plume from the fires, as well as the altitude as derived by the AER_LH product algorithm.

![Figure 140](image)

Figure 140: TROPOMI AER_LH product on 11 January 2020 over the south Pacific, showing the altitude as derived by the S5P AER_LH algorithm of the fires smoke from Australian bush fires.

The smoke provides an opportunity to compare the AER_LH with CALIOP measurements, since the extent of the smoke plume is so large that the CALIPSO satellite track intersects with the plume almost daily. An inspection of CALIOP quick-looks revealed much higher altitudes of the smoke derived by CALIOP than by TROPOMI.

A comparison of the CALIOP backscatter data and AER_LH data as before is presented below for 11 January 2020. In Figure 141 the AER_LH product for 11 January 2020 is plotted again over a VIIRS RGB picture, showing the smoke over clouds and in clear sky (the ALH is retrieved only in clear sky pixels). The maximum altitude in the AER_LH data is about 13 km. However, the CALIOP data show much higher altitudes for the plume. In Figure 142 the CALIOP total attenuated backscatter at 532 nm is shown for the yellow track shown in Figure 141. The plume can be seen around about 44°S and 110°E at an altitude between about 17 and 21 km, which is much higher than the S5P AER_LH. The AER_LH retrievals from TROPOMI are shown in the curtain plot as black and white dots as before. Clearly, the AER_LH is much lower than the altitude of the smoke plume.
Figure 141: NPP/VIIRS RGB image with S5P/TROPOMI AER_LH on 11 January over the south Pacific with the CALIPSO track of that day overlaid in yellow.

Figure 142: CALIOP aerosol backscatter rat 532 nm along the track shown in Figure 141 (eastern of Australia on 11 January 11, 2020).
The exact reason for the much lower altitude retrieved by the AER_LH algorithm is not clear, but it is obvious that altitudes above 20 km were not anticipated. The pressures at these altitudes are about 93 hPa (17 km) to 50 hPa (21 km) (Anderson et al., 1982). The AER_LH neural network (NN) was trained to perform within pressures of 1000-75 hPa, so the sensitivity of the algorithm to aerosols at this altitude is low at best. In the weeks after 11 January 2020 the plume kept clearly visible in CALIOP data and rose to even higher altitudes, up to even 30 km. At that altitude air pressures can be expected to be as low as 10 hPa. The AER_LH algorithm was not created to retrieve ALH at such low air pressures. A new NN may be trained to incorporate these extreme low air pressures. The need for such an extension will have to be investigated, as the occurrence of high altitude smoke like the case presented here seems rather rare. Furthermore, simulations will have to be performed first to test whether the AER_LH algorithm is at all sensitive to aerosols at such a high altitude, before this is to be included in the NN and operational algorithm.

Another issue that can play a role here is cloud contamination. As can be seen from Figure 141 and Figure 142, the area is very cloudy and the algorithm is known to be very sensitive to (residual) cloud contamination, and this will also bias the ALH low.
15 References

The validation activities and requirements applying to the operational phase of the S5P mission are described in the *S5P Cal/Val Plan for the Operational Phase* [S5P-CSCOP], the *S5P Geophysical Validation Requirements Document* [S5PVT-Req], the *Copernicus Sentinels 4 and 5 Mission Requirements Traceability Document* [S4/5-MRTD], and the recommendations formulated by ESL-L2 developers in their *Algorithm Theoretical Basis Documents* available on the ESA Copernicus Sentinel Online website [ER_CoperATBD].

15.1 Reference documents

[S5PVT-Req] Requirements for the Geophysical Validation of Sentinel-5 Precursor Products  
*source:* ESA; *ref:* S5P-RS-ESA-SY-164; *issue:* date: 2014-05-21

[S5P-CSCOP] ESA-EOPG-CSCOP-PL-0073, Sentinel-5 Precursor Calibration and Validation Plan for the Operational Phase  
*source:* ESA; *ref:* ESA-EOPG-CSCOP-PL; *issue:* 1; *revision:* 1; *date:* 2017-11-06

[S4/5-MRTD] Copernicus Sentinels 4 and 5 Mission Requirements Traceability Document  
*source:* ESA; *ref:* EOP-SM/2413/BV-bv; *issue:* 2; *revision:* 0; *date:* 2017-07-07


[JCGM-GUM] GUM: Joint Committee for Guides in Metrology (JCGM/WG 1) 100:2008, Evaluation of measurement data – Guide to the expression of uncertainty in a measurement (GUM)


[S5P-NomL1] Terms, definitions and abbreviations for TROPOMI L01b data processor;  
*source:* KNMI; *ref:* S5P-KNMI-L01B-0004-L1; *issue:* 3.0.0; *date:* 2013-11-08

[S5P-NomA] Terms and symbols in the TROPOMI Algorithm Team;  
*source:* KNMI; *ref:* SN-TROPOMI-KNMI-049; *issue:* 0.1.2; *date:* 2013-03-11

15.2 Peer-reviewed articles


15.3 Electronic references

[ER_TROPOMI] TROPOMI website [http://www.tropomi.eu]
[ER_VDAF] TROPOMI Validation Website / Validation Data Analysis Facility [http://mpc-vdaf.tropomi.eu]
[ER_L2QC] TROPOMI Portal for Level-2 Data Quality Control [http://mpc-l2.tropomi.eu]
[ER_S5PVVT] SSP Validation Team AO projects [https://earth.esa.int/web/guest/pi-community/apply-for-data/ao-s]
[ER_CoperESA] ESA Copernicus website [http://www.esa.int/copernicus]
[ER_C3S] Copernicus Climate Change Service (C3S) website [http://climate.copernicus.eu]
[ER_CODA] CODA Atmospheric Toolbox [https://atmospherictoolbox.org/coda]

ESA FRM Projects Websites

[ER_FRM4DOAS] Fiducial Reference Measurements for Ground-Based DOAS Air-Quality Observations project website [http://frm4doas.aeronomie.be]
[ER_Pandonia] Fiducial Reference Measurements for Ground-Based Direct-Sun Air-Quality Observations project [http://pandonia.net]
### Monitoring Networks Websites and Data Centres

| [ER_ACTRIS] | European Research Infrastructure for the observation of Aerosol, Clouds, and Trace gases website | [ER_Cloudnet] | Cloudnet remote sensing network website |
| [ER_COCCON] | Collaborative Carbon Column Observing Network (COCCON) website | [ER_EARLINET] | European Aerosol Research Lidar Network (EARLINET) website |
| [ER_EUBREWNET] | COST Action for a coherent network of European Brewer Spectrophotometer monitoring stations (EUBREWNET) website | [ER_EUMETNET] | European Meteorological Services Network (EUMETNET) website |
| [ER_EVDC] | ESA Validation Data Centre (EVDC) website | [ER_NDACC] | Network for the Detection of Atmospheric Composition Change (NDACC) website |
| [ER_NOVAC] | Network for Observation of Volcanic and Atmospheric Change (NOVAC) website | [ER_SHADOZ] | Southern Hemisphere ADditional OZonesonde programme website |
| [ER_TCCON] | Total Carbon Column Observing Network (TCCON) website | [ER_TOLnet] | Tropospheric Ozone Lidar Network (TOLnet) website |
| [ER_WOUDC] | World Ozone and Ultraviolet Data Centre (WOUDC) website | |

[http://www.actris.eu](http://www.actris.eu)  
[https://cloudnet.fmi.fi](https://cloudnet.fmi.fi)  
[https://www.imk-asf.kit.edu/english/COCCON.php](https://www.imk-asf.kit.edu/english/COCCON.php)  
[http://www.earlinet.org](http://www.earlinet.org)  
[http://www.eubrewnet.org](http://www.eubrewnet.org)  
[http://eumetnet.eu](http://eumetnet.eu)  
[http://evdc.esa.int](http://evdc.esa.int)  
[http://ndacc.org](http://ndacc.org)  
[https://tropo.gsfc.nasa.gov/shadoz](https://tropo.gsfc.nasa.gov/shadoz)  
[https://tccon-wiki.caltech.edu](https://tccon-wiki.caltech.edu)  
[http://www-air.larc.nasa.gov/missions/TOLNet](http://www-air.larc.nasa.gov/missions/TOLNet)  
[http://woudc.org](http://woudc.org)
16 Acknowledgements

This Section acknowledges the authors of this report in charge of the ATM-MPC S5P Routine Operations validation service (Table 23), the operators of S5P validation facilities, the providers of Fiducial Reference Measurements and other validation data, and the support provided by the Agencies.

16.1 ATM-MPC S5P Routine Operations Validation Service

Table 23 – Responsibilities for the ATM-MPC S5P routine operations validation service: Product Validation Coordinators responsible for validation and reporting per data product (third column), and Product Validation Contributors participating in the validation and reporting per data product (fourth column).

<table>
<thead>
<tr>
<th>S5P Product ID</th>
<th>Geophysical Quantity</th>
<th>Product Coordinator for Routine Operations Validation Activities</th>
<th>Product Contributors to Routine Operations Validation Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1B</td>
<td>Radiance and irradiance</td>
<td>A. Ludewig (KNMI)</td>
<td>M. Coldewey-Egbers (DLR)</td>
</tr>
<tr>
<td>L2_O3</td>
<td>O₃ total column</td>
<td>T. Verhoelst (BIRA-IASB)</td>
<td>K. Garane (AUTH)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>K.-P. Heue (DLR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C. Lerot (BIRA-IASB)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A. Redondas (AEMET)</td>
</tr>
<tr>
<td>L2_O3_PR</td>
<td>O₃ profile</td>
<td>A. Keppens (BIRA-IASB)</td>
<td>O. Tuinder (KNMI)</td>
</tr>
<tr>
<td>L2_O3_TCL</td>
<td>O₃ tropospheric column</td>
<td>D. Hubert (BIRA-IASB)</td>
<td>K.-P. Heue (DLR)</td>
</tr>
<tr>
<td>L2_NO2</td>
<td>NO₂ stratospheric column</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>K.-U. Eichmann (IUPB)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NO₂ tropospheric column</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NO₂ total column</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2_SO2</td>
<td>SO₂ total column</td>
<td>T. Wagner (MPI-C)</td>
<td>P. Hedelt (DLR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N. Theys (BIRA-IASB)</td>
</tr>
<tr>
<td>L2_HCHO</td>
<td>HCHO total column</td>
<td>K.-U. Eichmann (IUPB)</td>
<td>K.L. Chan (DLR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>S. Compernolle (BIRA-IASB)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>I. De Smedt (BIRA-IASB)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C. Viguouroux (BIRA-IASB)</td>
</tr>
<tr>
<td>L2_CO</td>
<td>CO total column</td>
<td>B. Langerock (BIRA-IASB)</td>
<td>M.K. Sha (BIRA-IASB)</td>
</tr>
<tr>
<td>L2_CH4</td>
<td>CH₄ total column</td>
<td>M.K. Sha (BIRA-IASB)</td>
<td>B. Langerock (BIRA-IASB)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A. Lorente Delgado (SRON)</td>
</tr>
<tr>
<td>L2_CLOUD</td>
<td>Cloud Fraction</td>
<td>S. Compernolle (BIRA-IASB)</td>
<td>R. Lutz (DLR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P. Wang (KNMI)</td>
</tr>
<tr>
<td></td>
<td>Cloud Height</td>
<td></td>
<td>A. Argyrouli (DLR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cloud Optical Thickness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2_AER_AI</td>
<td>Aerosol Absorbing Index</td>
<td>T. Wagner (MPI-C)</td>
<td>D. Stein Zweers (KNMI)</td>
</tr>
<tr>
<td>L2_AER_LH</td>
<td>Aerosol Layer Height</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>M. de Graaf (KNMI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>K. Michailidis (AUTH)</td>
</tr>
</tbody>
</table>
16.2 S5P validation facilities

The ATM-MPC Validation Data Analysis Facility (VDAF) hosted at BIRA-IASB by S. Compernolle, J.-C. Lambert and B. Langerock, runs the S5P TROPOMI Automated Validation Server (VDAF-AVS) [ER_VDAF-AVS] developed and operated jointly by s&t and at BIRA-IASB. The VDAF-AVS server is based on the HARP toolset developed and maintained by S. Niemeijer at s&t. The VDAF also hosts the dedicated S5P TROPOMI validation website [ER_VDAF].

Part of the validation work for trace gases data relies on the Multi-TASTE versatile validation system, developed and operated at BIRA-IASB by S. Compernolle, J. Granville, D. Hubert, A. Keppens, J.-C. Lambert, and T. Verhoelst. Multi-TASTE has been supported by the Belgian Federal Science Policy Office (BELSPO), with ad hoc support provided by the EC, ESA and EUMETSAT for specific satellite validation and metrology applications.

Part of the total ozone and aerosol validation work makes use of the validation facility operated at AUTH, and developed by D. Balis, K. Garane, ML. Koukouli and K. Michailidis with support from ESA and EUMETSAT.

The ESA Atmospheric Validation Data Centre (EVDC) [ER_EVDC], hosted at the Norwegian Institute for Air Research (NILU) under the supervision of A.M. Fjæraa, is acknowledged for facilitating access to the validation data from ground-based monitoring networks and field campaigns.

16.3 Validation data

The ground-based data used in this study was obtained as part of the Brewer and Dobson ozone column monitoring networks ([ER_WOUDC], [ER_EUBREWNET]), the Network for the Detection of Atmospheric Composition Change (NDACC) [ER_NDACC], Southern Hemisphere Additional Ozone programme (SHADOZ) [ER_SHADOZ], and the Total Carbon Column Observation Network (TCCON) [ER_TCCON], all contributors to WMO’s Global Atmosphere Watch (GAW). Data archived in the associated data centres and lists of associated data originators are publicly available.

Instrument PIs, the scientific teams and the staff at the stations are thanked warmly for special processing efforts and faster data delivery dedicated to TROPOMI validation:


- Rapid delivery O₃ profile data from the SHADOZ network was organised in the framework of the S5PVT AO project CHEOPS-5p (ID #28587, PIs A. Keppens and J.-C. Lambert, BIRA-IASB, Co-Is D. Balis, D. Hubert, W. Steinbrecht, T. Stavrakou, A. Delcloo, S. Godin-Beekmann, T. Leblanc, R. Stübi, A.M. Thompson, T. Verhoelst, G. Ancellet, and V. Duflot). Rapid delivery ozonesonde profile data were also provided by KNMI (A. Piters, M. Allaart) and NOAA (B.J. Johnson).
• Rapid delivery NO₂ data from NDACC MAX-DOAS and ZSL-DOAS stations was gathered in the framework of the S5PVT AO projects CESAR (ID #28596, PI A. Apituley, KNMI) and NIDFORVAL (ID #28607, PI G. Pinardi, BIRA-IASB). The LATMOS SAOZ_RT team (A. Pazmino, A. Bazureau, F. Goutail, and J.-P. Pommeneau) at IPSL/UVSQ/UPMC-CNRS is thanked for the near-real-time processing and delivery of ZSL-DOAS SAOZ data. ESA's FRM programme and LuftBlick/U. Innsbruck (A. Cede, M. Gebetsberger and M. Tiefengraber) are acknowledged for the rapid delivery of total NO₂ data from the Pandonia Global Network (PGN).

• Rapid delivery HCHO data from NDACC FTIR and MAX-DOAS stations was gathered in the framework of the S5PVT AO projects CESAR (ID #28596, PI A. Apituley, KNMI) and NIDFORVAL (ID #28607, Co-PIs G. Pinardi and C. Vigouroux, BIRA-IASB). This work could not be possible without the work of the FTIR and DOAS spectra and/or data providers: Carlos Augusto Bauer Aquino (IFRO); Cornelis Becker (SAHO); Thomas Blumenstock and Amelie Röhling (KIT, IMK-ASF); Martine De Mazière, Christian Hermans, François Hendrick, Michel Van Roozendael and Minqiang Zhou (BIRA); Omaira García (AEMET); Michel Grutter, Claudia Rivera, Alejandro Bezanilla, César Guarin and Wolfgang Stremme (UNAM); James Hannigan and Ivan Ortega (NCAR); Pascal Jesseck and Yao Té (LERMA-IPSL); Nicholas Jones and Clare Paton-Walsh (Univ. Wollongong), Rigel Kivi (FMI), Erik Lutsch and Kim Strong (Univ. Toronto); Maria Makarova and Anatoly Poberovskii (Univ. St. Petersburg); Emmanuel Mahieu and Christian Servais (Univ. Liège); Jean-Marc Metzger (Univ. Reunion Island); Isamu Morino and Hideaki Nakajima (NIES); Isao Murata (Univ. Tohoku); Tomoo Nagahama (ISEE); Justus Notholt, Mathias Palm and Holger Winkler (Univ. Bremen); Markus Rettinger and Ralf Sussman (KIT, IMK-IFU); John Robinson and Dan Smale (NIWA); Pucai Wang, Youwen Sun and Cheng Liu (CAS); Ankie Piters (KNMI); Thomas Wagner, Sebastian Donner and Julia Remmers (MPIC).

• Rapid delivery of CO and CH₄ data from COCCON FTIR stations was gathered in the framework of the ESA projects COCCON-PROCEEDS and Sentinel-5p MPC-VDAF.

• Rapid delivery of CO and CH₄ data from TCCON FTIR stations was gathered in the framework of the S5PVT AO project TCCON4S5P (ID #28603, PI M. Kumar Sha, BIRA-IASB).

• Rapid delivery of NDACC data is partly supported by the CAMS-27 data procurement service contract by ECMWF for the validation of the Copernicus Atmospheric Monitoring Service (CAMS).

CLOUDNET classification product was obtained via the European Research Infrastructure for the observation of Aerosol, Clouds, and Trace gases (ACTRIS) [ER_ACTRIS] and EVDC. Data was processed at the Department of Meteorology, University of Reading, UK, and at the Finnish Meteorological Institute. They acknowledge funding from the EU’s Horizon 2020 programme under grant agreement No 654109 and the Cloudnet project (EU contract EVK2-2000-00611).

Automated Lidars and Ceilometers (ALC) data was obtained as part of the E-PROFILE observation programme run in the framework of the European Meteorological Services Network (EUMETNET) [ER_EUMETNET].

EUMETSAT AC-SAF and DLR are acknowledged for the provision of Metop-A and Metop-B GOME-2 ozone and cloud data.
EARLINET is acknowledged for providing aerosol LIDAR profiles available at https://data.earlinet.org/. The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654109 and previously from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n°262254. KNMI is acknowledged for the provision of EOS-Aura OMI O$_3$, NO$_2$, HCHO and UVAI data. The OMI QA4ECV data records are an outcome of the EC FP7-SPACE-2013-1 project No 607405: Quality Assurance for Essential Climate Variables (QA4ECV).

NASA/GSFC is acknowledged for the provision of (i) Suomi-NPP OMPS radiance, O$_3$ and UVAI data, (ii) Suomi-NPP VIIRS cloud data obtained with a pre-production code run specifically for limited S5P team analysis, (iii) EOS-Aqua MODIS cloud fraction, cloud top height and cloud optical thickness data, and (iv) MISR and CALIOP aerosol layer height data.

16.4 Agency support

The ATM-MPC S5P routine operations validation service is supported jointly by ESA, the Belgian Federal Science Policy Office (BELSPO) through BIRA-IASB, the Netherlands Space Office (NSO), and the German Aerospace Centre (DLR). S5PVT Announcement of Opportunity (AO) projects [ER_S5PVT] having contributed to this report are funded by several national agencies from Europe, Canada, China, Japan and the USA.
17 Terms, definitions and abbreviated terms

17.1 Terms and definitions

accuracy  closeness of agreement between a quantity value obtained by measurement and the true value of the measurand; note that it is not a quantity and it is not given a numerical quantity value [JCGM-VIM]

bias  (1) systematic error of indication of a measuring system [JCGM-VIM]
       (2) estimate of a systematic measurement error [JCGM-VIM]

error  (1) measured quantity value minus a reference quantity value [JCGM-VIM]
       (2) difference of quantity value obtained by measurement and true value of the measurand (CEOS/ISO)

influence quantity  quantity that, in a direct measurement, does not affect the quantity that is actually measured, but affects the relation between the indication and the measurement result [JCGM-VIM]

Level-1b data  calibrated, geo-located Earth reflectance and radiance spectra in all spectral bands; solar irradiance data, annotation data and references to used calibration data

Level-2 data  geophysical measurand at the same resolution and geolocation as the Level 1 source data from which it is derived

Level-3 data  data or retrieved geophysical parameters (i.e. derived from Level 1 or 2 data products) mapped on uniform space-time grid scales, usually with some completeness and consistency. Such re-sampling may include averaging, compositing, kriging, use of Kalman filters…

measurand  quantity intended to be measured [JCGM-VIM]

measurement bias  estimate of a systematic measurement error [JCGM-VIM]

measurement error  measured quantity value minus a reference quantity value [JCGM-VIM]

uncertainty  non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used [JCGM-VIM]

precision  closeness of agreement between quantity values obtained by replicate measurements of a quantity on the same or similar object under specified conditions [JCGM-VIM]

random error  component of measurement error that in replicate measurements varies in an unpredictable manner; note that random measurement error equals measurement error minus systematic measurement error [JCGM-VIM]

relative standard uncertainty  standard measurement uncertainty divided by the absolute value of the measured quantity value [JCGM-VIM]

stability  ability of a measuring system to maintain its metrological characteristics constant with time [JCGM-VIM]

systematic error  component of measurement error that in replicate measurements remains constant or varies in a predictable manner [JCGM-VIM]

uncertainty  non-negative parameter that characterizes the dispersion of the quantity values that are being attributed to a measurand, based on the information used [JCGM-VIM]

validation  (1) the process of assessing, by independent means, the quality of the data products derived from the system outputs (CEOS/ISO)
            (2) verification where the specified requirements are adequate for an intended use [JCGM-VIM]

verification  the provision of objective evidence that a given data product fulfils specified requirements; note that, when applicable, measurement uncertainty should be taken into consideration [JCGM-VIM]
17.2 Acronyms and abbreviations

A(A)I  Aerosol (Absorbing) Index
AC-SAF Atmospheric Composition Satellite Application Facility
ACTRIS European Research Infrastructure for the observation of Aerosol, Clouds, and Trace gases
AK  averaging kernel
ALC  Automated Lidars and Ceilometers network
AMF  Air Mass Factor
AO  Announcement of Opportunity
ARM  Atmospheric Radiation Measurement program
ATBD Algorithm Theoretical Basis Document
AVS  Automated Validation Server
AUTH  Aristotle University of Thessaloniki
BELSPO  Belgian Federal Science Policy Office
BIRA-IASB  Royal Belgian Institute for Space Aeronomy
C3S  Copernicus Climate Change Service
CAL  Clouds As Layers
CAMS  Copernicus Atmosphere Monitoring Service
CCD  Convective Cloud Differential method
CCI  Climate Change Initiative
CESAR  Cabauw Experimental Research Site for Atmospheric Research
CF  Cloud Fraction (fractional cloud cover)
CHEOPS-5p Validation of Copernicus HEight-resolved Ozone data Products from Sentinel-5p
CLOUDNET  Cloud properties monitoring Network
COCCON  Collaborative Carbon Column Observing Network
COT  Cloud Optical thickness
CRB  Clouds as Reflecting Boundaries
CRG  Climate Research Group
C(T)H  Cloud (Top) Height
DFS  Degree of Freedom of the Signal
DLR  German Aerospace Centre / Deutsches Zentrum für Luft- und Raumfahrt
DOAS  Differential Optical Absorption Spectroscopy
DU  Dobson Unit
EARLINET  European Aerosol Research Lidar Network
EC  European Commission
ECMWF  European Centre for Medium-Range Weather Forecasts
EOS  Earth Observing System
EPS  EUMETSAT Polar System
ESA  European Space Agency
ESL  Expert Support Laboratory
EU  European Union
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUMETNET</td>
<td>European Meteorological Services Network</td>
</tr>
<tr>
<td>EUMETSAT</td>
<td>European Organisation for the Exploitation of Meteorological Satellites</td>
</tr>
<tr>
<td>EVDC</td>
<td>ESA Atmospheric Validation Data Centre</td>
</tr>
<tr>
<td>FRM</td>
<td>Fiducial Reference Measurement</td>
</tr>
<tr>
<td>FTIR</td>
<td>Fourier Transform Infra-Red</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full Width at Half Maximum</td>
</tr>
<tr>
<td>GE_LER</td>
<td>Geometry-dependent Effective Lambert-Equivalent Reflectivity</td>
</tr>
<tr>
<td>GAW</td>
<td>Global Atmosphere Watch</td>
</tr>
<tr>
<td>GOME(-2)</td>
<td>Global Ozone Monitoring Experiment(-2)</td>
</tr>
<tr>
<td>GOSAT(-2)</td>
<td>Greenhouse gases Observing Satellite(-2)</td>
</tr>
<tr>
<td>GSFC</td>
<td>Goddard Space Flight Center</td>
</tr>
<tr>
<td>GUM</td>
<td>Guide to the Expression of Uncertainty in Measurement</td>
</tr>
<tr>
<td>IPSL/UVSQ</td>
<td>Institut Pierre-Simon Laplace / Université de Versailles Saint-Quentin-en-Yvelines</td>
</tr>
<tr>
<td>IUP-UB</td>
<td>Institute of Environmental Physics - University of Bremen</td>
</tr>
<tr>
<td>JCGM</td>
<td>Joint Committee for Guides in Metrology</td>
</tr>
<tr>
<td>KNMI</td>
<td>Koninklijk Netherlands Meteorologisch Instituut / Royal Dutch Meteorological Institute</td>
</tr>
<tr>
<td>LATMOS</td>
<td>Laboratoire Atmosphères, Milieux, Observations Spatiales</td>
</tr>
<tr>
<td>LER</td>
<td>Lambert-equivalent reflectivity</td>
</tr>
<tr>
<td>Lidar</td>
<td>Light Detection And Ranging</td>
</tr>
<tr>
<td>MAX-DOAS</td>
<td>Multi Axis Differential Optical Absorption Spectroscopy</td>
</tr>
<tr>
<td>Metop</td>
<td>polar orbiting Meteorological Operational satellite</td>
</tr>
<tr>
<td>MPC</td>
<td>Mission Performance Centre</td>
</tr>
<tr>
<td>MPI-C</td>
<td>Max Planck Institute for Chemistry</td>
</tr>
<tr>
<td>NASA</td>
<td>National Aeronautics and Space Administration</td>
</tr>
<tr>
<td>NDACC</td>
<td>Network for the Detection of Atmospheric Composition Change</td>
</tr>
<tr>
<td>NIDFORTVAL</td>
<td>S5P Nitrogen Dioxide and FORmaldehyde VALidation using NDACC and complementary FTIR and UV-Vis DOAS ground-based remote sensing data</td>
</tr>
<tr>
<td>NOVAC</td>
<td>Network for Observation of Volcanic and Atmospheric Change</td>
</tr>
<tr>
<td>NILU</td>
<td>Norsk Institutt for Luftforskning / Norwegian Institute for Air Research</td>
</tr>
<tr>
<td>NOAA</td>
<td>National Oceanic and Atmospheric Administration</td>
</tr>
<tr>
<td>NRT</td>
<td>Near Real Time</td>
</tr>
<tr>
<td>NSO</td>
<td>Netherlands Space Office</td>
</tr>
<tr>
<td>PANDORA</td>
<td>not an acronym; direct Sun UV-visible spectrometer</td>
</tr>
<tr>
<td>OFFL</td>
<td>Off-line</td>
</tr>
<tr>
<td>OMI</td>
<td>Ozone Monitoring Instrument</td>
</tr>
<tr>
<td>OMPS</td>
<td>Ozone Mapper and Profiling Suite</td>
</tr>
<tr>
<td>PDGS</td>
<td>Payload Data Ground Segment</td>
</tr>
<tr>
<td>PI</td>
<td>Principal Investigator</td>
</tr>
<tr>
<td>PICS</td>
<td>Pseudo-Invariant Calibration Site</td>
</tr>
<tr>
<td>PRF</td>
<td>Product Readme File</td>
</tr>
<tr>
<td>PUM</td>
<td>Product User Manual</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>QA4EO</td>
<td>Quality Assurance framework for Earth Observation</td>
</tr>
<tr>
<td>QC</td>
<td>Quality Control</td>
</tr>
<tr>
<td>QWG</td>
<td>Quality Working Group</td>
</tr>
<tr>
<td>RAL</td>
<td>Rutherford Appleton Laboratory</td>
</tr>
<tr>
<td>SSP</td>
<td>Sentinel-5 Precursor</td>
</tr>
<tr>
<td>SSPVT</td>
<td>Sentinel-5 Precursor Validation Team</td>
</tr>
<tr>
<td>SAOZ</td>
<td>Système d’Analyse par Observation Zénithale</td>
</tr>
<tr>
<td>SCIAMACHY</td>
<td>SCanning Imaging Absorption spectroMeter for Atmospheric CartograpHY</td>
</tr>
<tr>
<td>SHADOZ</td>
<td>Southern Hemisphere ADditional O Zonesonde programme</td>
</tr>
<tr>
<td>SRON</td>
<td>Netherlands Institute for Space Research</td>
</tr>
<tr>
<td>Suomi-NPP</td>
<td>Suomi National Polar-orbiting Partnership</td>
</tr>
<tr>
<td>TCCON</td>
<td>Total Carbon Column Observing Network</td>
</tr>
<tr>
<td>TCCON4S5P</td>
<td>Validation of S5P Methane and Carbon Monoxide with TCCON Data</td>
</tr>
<tr>
<td>TOLNet</td>
<td>Tropospheric Ozone Lidar Network</td>
</tr>
<tr>
<td>TROPOMI</td>
<td>Tropospheric Monitoring Instrument</td>
</tr>
<tr>
<td>UTLS</td>
<td>Upper Troposphere / Lower Stratosphere</td>
</tr>
<tr>
<td>UVAI</td>
<td>Ultraviolet aerosol absorbing index</td>
</tr>
<tr>
<td>VDAF</td>
<td>Validation Data Analysis Facility</td>
</tr>
<tr>
<td>VIIRS</td>
<td>Visible Infrared Imaging Radiometer Suite</td>
</tr>
<tr>
<td>VIM</td>
<td>International Vocabulary of Metrology</td>
</tr>
<tr>
<td>WMO</td>
<td>World Meteorological Organization</td>
</tr>
<tr>
<td>WOUDC</td>
<td>World Ozone and Ultraviolet Data Centre</td>
</tr>
<tr>
<td>ZSL-DOAS</td>
<td>Zenith-Scattered-Light Differential Optical Absorption Spectroscopy</td>
</tr>
</tbody>
</table>

**END OF DOCUMENT**