Sentinel-5 Precursor
Mission Performance Centre

Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products #11: April 2018 – June 2021

Prepared by	Sentinel-5 Precursor Mission Performance Centre
Reference	S5P-MPC-IASB-ROCVR-11.00.03-20210609
CI identification	DI-MPC-ROCVR
Document update	#11
Issue	11.00.03
Date of issue	2021-06-09
Status	FINAL DRAFT
Distribution	public

Approval Record		
Checked by:	J.-C. Lambert (BIRA-IASB) Q. L. Kleipool (KNMI) D. Loyola (DLR) J. P. Veefkind (KNMI) A. Dehn (ESA)	MPC ESL-VAL Lead MPC ESL-L1 Lead MPC ESL-L2 Lead MPC Technical Manager MPC Project Officer
Approved by:	A. Dehn (ESA)	ESA Data Quality Manager
Signatures:	J.-C. Lambert (BIRA-IASB) A. Dehn (ESA)	

Document Information

Document Identification	
Title	Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products #11: April 2018 – June 2021
Type of document	S5P MPC Routine Operations Consolidated Validation Report (ROCVR)
Document ID	S5P-MPC-IASB-ROCVR-11.00.03-20210609
ROCVR update Issue number Date of issue Status Distribution	#11 version 11.00.03 2021-06-09 FINAL DRAFT public
Available on:	https://mpc-vdaf.tropomi.eu
Editors	J.-C. Lambert and A. Keppens (BIRA-IASB)
Contributors	S. Compernolle (BIRA-IASB), K.-U. Eichmann (IUP-UB), M. de Graaf (KNMI), D. Hubert (BIRA-IASB), A. Keppens (BIRA-IASB), Q. Kleipool (KNMI), B. Langerock (BIRA-IASB), M.K. Sha (BIRA-IASB), T. Verhoelst (BIRA-IASB), T. Wagner (MPI-C), C. Ahn (NASA/GSFC), A. Argyrouli (DLR), D. Balis (AUTH), K.L. Chan (DLR), I. De Smedt (BIRA-IASB), H. Eskes (KNMI), A.M. Fjæraa (NILU), K. Garane (AUTH), J.F. Gleason (NASA/GSFC), F. Goutail (LATMOS), J. Granville (BIRA-IASB), P. Hedelt (DLR), K.-P. Heue (DLR), G. Jaross (NASA/GSFC), ML. Koukouli (AUTH), J. Landgraf (SRON), R. Lutz (DLR), S. Nanda (KNMI), S. Niemeijer (s[&]t), A. Pazmiño (LATMOS), G. Pinardi (BIRA-IASB), J.-P. Pommereau (LATMOS), A. Richter (IUP-UB), N. Rozemeijer (KNMI), M. Sneep (KNMI), D. Stein Zweers (KNMI), N. Theys (BIRA-IASB), G. Tilstra (KNMI), O. Torres (NASA/GSFC), P. Valks (DLR), J. van Geffen (KNMI), C. Vigouroux (BIRA-IASB), P. Wang (KNMI), and M. Weber (IUP-UB)
Citation	Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products #11: April 2018 – June 2021. Lambert, J.-C., S. Compernolle, K.-U. Eichmann, M. de Graaf, D. Hubert, A. Keppens, Q. Kleipool, B. Langerock, M.K. Sha, T. Verhoelst, T. Wagner, C. Ahn, A. Argyrouli, D. Balis, K.L. Chan, I. De Smedt, H. Eskes, A.M. Fjæraa, K. Garane, J.F. Gleason, F. Goutail, J. Granville, P. Hedelt, K.-P. Heue, G. Jaross, ML. Koukouli, J. Landgraf, R. Lutz, S. Nanda, S. Niemeijer, A. Pazmiño, G. Pinardi, J.-P. Pommereau, A. Richter, N. Rozemeijer, M. Sneep, D. Stein Zweers, N. Theys, G. Tilstra, O. Torres, P. Valks, J. van Geffen, C. Vigouroux, P. Wang, and M. Weber. S5P MPC Routine Operations Consolidated Validation Report series, Issue #11, Version 11.00.03, 171 pp., April 17, 2021.

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 S5P Mission Performance Centre (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) AO projects [ER_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 till June 2021. It includes validation results for the UPAS processor upgrade to version 02.01.04 operational since November/December 2020, and validation results for the NLL2DP processor upgrade to version 01.04.00 operational since early December 2020.

Radiance and Irradiance

The validation of the wavelength assignment of the S5P L1B_UVN v01.00.00 products concludes to an agreement of 0.02 to 0.04 nm, which is within the pre-launch calibration accuracy. Initial validation of the L1B_RA reflectance with respect to OMI and OMPS independent satellite data indicates that TROPOMI is within 5% for the shorter wavelengths in band 3 and improving to 2% towards the longer wavelengths in band 4. For the short wave UV in band 1 TROPOMI L1B_RA is within 8% +/-2% of the expected modelled reflectance. In general, radiometric errors in bands 1 and 3 are large but they vary slowly over wavelength and most L2 retrievals are insensitive to such errors. Additional validation indicates that for bands 3 to 7 the mission requirements for the reflectance are met. The largest source of error in the reflectance is due to the initial pre-launch irradiance calibration. This is a known issue that will be addressed in future updates of the L1B data products.

The validation of the TROPOMI L1B_IR irradiance product shows that it is within 3 to 10% depending on the used reference spectrum, and also that there is a radiometric mismatch between band 2 and 3. Additional validation with other solar irradiance spectra shows that the difference exhibits smooth wavelength dependence, most likely caused by optical setup effects during the on-ground calibration. This anomaly affects the UV and UVIS channels, and can be corrected for in the update of the Level-0-to-1b processor. After this correction the difference with respect to reference spectra reduces to 2% and is within the expected radiometric accuracy.

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 also 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 3.5% 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 02.01.03 on 16 July 2020 (NRTI) and 13 July 2020 (OFFL), and to version 02.01.04 in November 2020, improves the agreement of the NRTI product with ground-based measurements at sites with complex surface albedo (snow/ice).

The comparison of S5P TROPOMI 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 +16% or +3.1 DU) and the mean dispersion of the differences (about 25% 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-2 processor was upgraded to version 02.01.03 in July 2020, and to version 02.01.04 in November 2020. Comparisons to ozonesonde do not indicate a clear change in data quality between Level-2 versions 01.x and 02.x.

The bias between S5P and other data sets exhibits a seasonal cycle, which is most easily visible from satellite intercomparisons. A convincing pattern of more elevated positive biases (5 DU or 25%) during the biomass burning season emerges at sites over the northern part of the Atlantic equatorial basin. However, at Atlantic sites in the Southern Hemisphere there is no consistent annual repetition of such a pattern. 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.

Nitrogen Dioxide

The three S5P L2_NO2 data products (tropospheric, stratospheric, and total column) up to version 01.04.00 (NRTI, OFFL) and 01.02.02 (RPRO) are in good overall agreement with correlative ground-based measurements. This assessment covers S5P operation for the period from April 2018 until May 2021. Correlative measurements from the MAX-DOAS network, the NDACC ZSL-DOAS/SAOZ network, and the Pandonia Global Network (PGN), respectively, as well as corresponding satellite data products (OMI), are used as reference for the validation. The bias between S5P and ground-based data varies with the total column of NO_2 except for the stratospheric column. Ground-based validation concludes to rather similar bias and uncertainty estimates for the L2_NO2 NRTI and OFFL/RPRO datasets.

The L2_NO2 tropospheric column data is compared to MAX-DOAS ground-based data from 25 stations. A negative median bias of -34% is found which depends on pollution level, being positive (11%) over clean areas (<2 Pmolec/cm^2) and negative (-46%) over highly polluted areas (> 15 Pmolec/cm^2). The bias is within the mission requirement of 50%. This bias estimate is reduced by up to 20% when the MAX-DOAS profile data are vertically smoothed using the S5P averaging kernels. The median dispersion of about 2.7 Pmolec/cm^2 exceeds the mission precision requirements (0.7 Pmolec/cm^2). It is only within limits for very clean sites. Comparison of TROPOMI tropospheric NO_2 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 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 NO_2 in the zenith-scattered-light (ZSL) geometry and for the NO_2 diurnal cycle, S5P reports stratospheric NO_2 column values generally lower by approximately 0.17 Pmolec/cm^2. The median bias of -6.1% is within the S5P mission requirements (0.2-0.4 Pmolec/cm^2) as well as the dispersion of 0.3 Pmolec/cm^2, considering the combined random errors and irreducible co-location mismatches. S5P comparisons to PGN total NO_2 column data at 3 high altitude sites conclude to a bias of approximatively -10%.

The L2_NO2 total NO_2 column data are compared to ground-based Pandora column data at 35 PGN sites. The median bias between S5P and PGN data is -7% with a dispersion of 1.5 Pmolec/cm^2. Comparison results vary with the total amount of NO_2, concluding to a positive bias of 4.1% over clean sites and high mountain areas (< 6 Pmolec/cm^2) and a negative bias of about -23% over polluted sites.

The version upgrade to 01.04.00 improved both bias and dispersion for the tropospheric and total columns. Comparing 3-month periods (JFM) of different years, a bias reduction of 33% for tropospheric NO_2 and 60% total NO_2 was found, respectively.

Formaldehyde

The S5P L2_HCHO column data up to version 02.01.04 (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. The processor upgrade to version 02.01.04 in December 2020 had no impact on data quality.

Comparisons to data from 28 FTIR stations conclude to a negative bias of -29% for high emission sites (> 8 Pmolec/cm^2) and a positive bias of 27% for clean sites (< 2.5 Pmolec/cm^2) when vertical smoothing differences are minimized by application of the averaging kernels. The bias with respect to NDACC MAX-DOAS HCHO measurements is slightly higher, about -37%. This bias reduces to about -20% when accounting for the difference in vertical smoothing between TROPOMI and MAX-DOAS data (using averaging kernels). All biases are within the S5P mission requirements (40-80%).

The dispersion of about 8 Pmolec/cm^2 (vs. NDACC FTIR data at clean sites) and 10 Pmolec/cm^2 (vs. MAX-DOAS) is within the S5P uncertainty mission requirements of 12 Pmolec/cm^2. These values are based on the use of median deviations to reduce the influence of large 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/cm^2 when considering regionally averaged columns.

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 S5P 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 SO_2 data product before and after the processor switch.

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 and TCCON 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 and TCCON, which is also within the mission requirement for precision (better than 10%). The averaged correlation coefficient reaches 0.9 for both NDACC and TCCON. From July 3 2019, onwards the NRTI processor uses the same settings as the OFFL processor and both products perform similarly since then.

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 and TCCON FTIR monitoring networks. The standard and bias-corrected S5P xCH4 column data exhibit a negative bias against TCCON of -0.68% and -0.25% respectively, which falls well within the mission requirement (bias of maximum 1.5%). The standard deviation of the relative bias 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. Since March 11 2020, 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.

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. Comparing spatially co-located measurements, the CLOUD CRB cloud fraction is slightly below that of S5P FRESCO, with higher deviations at high-latitude and/or coastal sites. 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.

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, deviations at high-latitude and/or coastal sites are reduced (e.g., Ny-Ålesund, Iquique). 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 dispersion between S5P and other cloud height data exceeds the S5P mission requirements, but likely the comparison error (due to the different sensitivity) is an important cause of this apparent discrepancy.

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 also a North-South gradient for cloud albedo. These patterns are reduced for processor versions 02.01.03 and 02.01.04.

Aerosol Index

The S5P L2_AER_AI (NRTI and OFFL) UV Aerosol Absorbing Index data 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 is currently larger than 1 UVAI unit as compared to OMI and OMPS. The reasons for this increasing bias are related to wavelength-dependent degradation and corrections are under investigation.

Aerosol Layer Height

The S5P 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). S5P 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.

The S5P 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 S5P 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 removes currently a large number of observations, which are in principle well suited for ALH retrievals.

Processing Baseline Identification

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

Product ID	Stream	Version	In operation from (orbit #, date)	In operation until (orbit #, date)
L1B_RA1/2/…/8		01.00.00	2818, 2018-04-30	current version
L1B_IR_UVN/SIR		01.00.00	2818, 2018-04-30	current version
L2_O3	NRTI	01.00.00 01.01.02 01.01.05 01.01.06 01.01.07 01.01.08 02.01.03 02.01.04	2955, 2018-05-09 4245, 2018-08-08 5931, 2018-12-05 7631, 2019-04-04 7999, 2019-04-30 12482, 2020-03-11 14285, 2020-07-16 16259, 2020-12-02	3943, 2018-07-18 5930, 2018-12-05 7631, 2019-04-04 7999, 2019-04-30 12482, 2020-03-11 14285, 2020-07-16 16256, 2020-12-02 current version
	OFFL	01.01.07 01.01.08 02.01.03 02.01.04	7907, 2019-04-23 12431, 2020-03-07 14239, 2020-07-13 16213, 2020-11-29	12431, 2020-03-07 14238, 2020-07-12 16212, 2020-11-29 current version
	RPRO	01.01.07	2818, 2018-04-30	7906, 2019-04-23
L2_O3_TCL	OFFL (CCD)	01.01.05 01.01.06 01.01.07 01.01.08 02.01.03 02.01.04	2824, 2018-04-30 7435, 2019-03-21 7804, 2019-04-16 12288, 2020-02-26 14132, 2020-07-02 16104, 2020-11-18	7421, 2019-03-20 7791, 2019-04-15 12273, 2020-02-25 14075, 2020-07-01 16090, 2020-11-17 current version
L2_NO2	NRTI	01.00.01 01.00.02 01.01.00 01.02.00 01.02.02 01.03.00 01.03.01 01.03.02 01.04.00	2955, 2018-05-09 3745, 2018-07-04 3947, 2018-07-18 5336, 2018-10-24 5931, 2018-12-05 7519, 2019-03-27 8000, 2019-04-30 8907, 2019-07-03 16259, 2020-12-02	3364, 2018-06-07 3946, 2018-07-18 5333, 2018-07-24 5929, 2018-12-05 7517, 2019-03-27 7999, 2019-03-30 8906, 2019-07-03 16256, 2020-12-02 current version
	RPRO	*01.02.02*	*2836, 2018-05-01*	*5235, 2018-10-17*
	OFFL	01.02.00 01.02.02 01.03.00 01.03.01 01.03.02 01.04.00	5236, 2018-10-17 5840, 2018-11-29 7425, 2019-03-20 7907, 2019-04-23 8815, 2019-06-26 16213, 2020-11-29	5832, 2018-11-28 7424, 2019-03-20 7906, 2019-04-23 8814, 2019-06-26 16212, 2020-11-29 current version
L2_HCHO	NRTI	01.00.00 01.01.01 01.01.02 01.01.05 01.01.06 01.01.07 01.01.08 02.01.03 02.01.04	2955, 2018-05-09 3947, 2018-07-18 4245, 2018-08-08 5931, 2018-12-05 7632, 2019-04-04 8000, 2019-04-30 12483, 2020-03-11 14285, 2020-07-16 16259, 2020-12-02	3943, 2018-07-18 4244, 2018-08-08 5929, 2018-12-05 7631, 2019-04-04 7999, 2019-04-30 12482, 2020-03-11 14285, 2020-07-16 16256, 2020-12-02 current version
	RPRO	*01.01.02* *01.01.05*	*2818, 2018-04-30* *3627, 2018-06-26*	*4147, 2018-08-01* *5832, 2018-11-28*
	OFFL	01.01.05 01.01.06 01.01.07	5832, 2018-11-27 7546, 2019-03-29 7907, 2019-04-23	7545, 2019-03-29 7906, 2019-04-23 12431, 2020-03-07

Product ID	Stream	Version	In operation from (orbit #, date)	In operation until (orbit #, date)
		01.01.08 02.01.03 02.01.04	12432, 2020-03-07 14239, 2020-07-13 16213, 2020-11-29	14238, 2020-07-12 16212, 2020-11-29 current version
L2_SO2	NRTI	01.00.00 01.01.01 01.01.02 01.01.07 01.01.08 02.01.03 02.01.04	2955, 2018-05-09 3953, 2018-07-18 4245, 2018-08-08 5931, 2018-12-05 12482, 2020-03-11 14285, 2020-07-16 16259, 2020-12-02	3939, 2018-07-17 4244, 2018-08-08 5929, 2018-12-05 12482, 2020-03-11 14285, 2020-07-16 16258, 2020-12-02 current version
	OFFL	01.00.00 01.01.01 01.01.02 01.01.07 01.01.08 02.01.03 02.01.04	3202, 2018-05-27 3848, 2018-07-10 4147, 2018-07-31 5933, 2018-11-28 12482, 2020-03-11 14239, 2020-07-13 16213, 2020-11-29	3847, 2018-07-10 4146, 2018-07-31 5932, 2018-11-28 12482, 2020-03-11 14238, 2020-07-12 16212, 2020-11-29 current version
L2_CO	NRTI	01.02.00 01.02.02 01.03.00 01.03.01 01.03.02 01.04.00	5336, 2018-10-24 5931, 2018-12-05 7519, 2019-03-27 7999, 2019-04-30 8906, 2019-07-03 16259, 2020-12-02	5929, 2018-12-05 7517, 2019-03-27 7999, 2019-04-30 8906, 2019-07-03 16256, 2020-12-02 current version
	RPRO[2]	01.02.02 01.03.01 01.03.02	0657, 2017-11-28 2818, 2018-04-30 2463, 2018-04-04	5346, 2018-10-25 5832, 2018-11-28 2477, 2018-04-05
	OFFL	01.02.02 01.03.00 01.03.01 01.03.02 01.04.00	5833, 2018-11-28 7425, 2019-03-20 7907, 2019-04-23 8815, 2019-06-26 16213, 2020-11-29	7424, 2019-03-20 7906, 2019-04-23 8814, 2019-06-26 16212, 2020-11-29 current version
L2_CH4	RPRO[2]	01.02.02 01.03.01 01.03.02	0657, 2017-11-28 2818, 2018-04-30 2463, 2018-04-04	5346, 2018-10-25 5832, 2018-11-28 2477, 2018-04-05
	OFFL	01.02.02 01.03.00 01.03.01 01.03.02 01.04.00	5833, 2018-11-28 7425, 2019-03-20 7907, 2019-04-23 8812, 2019-06-26 16213, 2020-11-29	7424, 2019-03-20 7906, 2019-04-23 8814, 2019-06-26 16212, 2020-11-29 current version
L2_CLOUD	NRTI	01.01.01 01.01.02 01.01.05 01.01.06 01.01.07 01.01.08 02.01.03 02.01.04	3947, 2018-07-18 4245, 2018-08-08 5931, 2018-12-05 7631, 2019-04-04 7999, 2019-04-30 12482, 2020-03-11 14285, 2020-07-16 16259, 2020-12-02	4244, 2018-08-08 5929, 2018-12-05 7631, 2019-04-04 7999, 2019-04-30 12482, 2020-03-11 14285, 2020-07-16 16256, 2020-12-02 current version
	OFFL	01.01.07 01.01.08 02.01.03 02.01.04	7907, 2019-04-23 12432, 2020-03-07 14239, 2020-07-13 16213, 2020-11-29	12431, 2020-03-07 14238, 2020-07-13 16212, 2020-11-29 current version
	RPRO	01.01.07	2818, 2018-04-30	7906, 2019-04-23

Product ID	Stream	Version	In operation from (orbit #, date)	In operation until (orbit #, date)
L2_AER_AI	NRTI	01.02.00 01.02.02 01.03.00 01.03.01 01.03.02 01.04.00	5336, 2018-10-24 5932, 2018-12-05 7519, 2019-03-27 8000, 2019-04-30 8906, 2019-07-03 16259, 2020-12-02	5929, 2018-12-05 7518, 2019-03-27 7999, 2019-04-30 8906, 2019-07-03 16258, 2020-12-02 current version
	OFFL	01.02.00 01.02.02 01.03.00 01.03.01 01.03.02 01.04.00	5236, 2018-10-17 5833, 2018-11-28 7425, 2019-03-20 7907, 2019-04-23 8815, 2019-06-26 16213, 2020-11-29	5832, 2018-11-28 7424, 2019-03-20 7906, 2019-04-23 8814, 2019-06-26 16212, 2020-11-29 current version
L2_AER_LH	NRTI	01.04.00	17155, 2021-01-03	current version
	OFFL	01.03.00 01.03.01 01.03.02 01.04.00	7425, 2019-03-20 7907, 2019-04-23 8815, 2019-06-26 16213, 2020-11-29	7906, 2019-04-23 8814, 2019-06-26 16212, 2020-11-29 current version
	RPRO	01.03.01	2818, 2018-04-30	7424, 2019-03-20

Table 1 – S5P TROPOMI data products and processor versions (NRTI near-real-time and OFFL off-line). Note 1: the operational phase (E2) of the S5P 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, dispersion and special features) have been derived for the following S5P operational data products on the basis of the validation results reported in this document:

Product ID	Stream	Product	Bias	Dispersion	Special features
L2_O3	NRTI	O_3 column	0.8%	2.5%	Larger dispersion over snow/ice due to coarse surface albedo climatology (up to but excluding v02.01.xx which has a dynamic surface albedo determination).
	OFFL	O_3 column	0.3%	2%	
L2_O3_TCL	OFFL (CCD)	O_3 tropospheric column	+16%	25%	Geographical imprints of sampling-related biases. Seasonal change of the bias. More elevated positive bias during biomass burning conditions in Northern Hemisphere.
L2_NO2	NRTI	NO_2 troposphere NO_2 stratosphere NO_2 total	-37% -5% 0±50%	2.7 Pmolec/cm² 0.3 Pmolec/cm² -	Bias (dispersion) depends on column amount. Troposphere: low pollution (<2 Pmolec/cm²) +10% (0.7 Pmolec/cm²); high pollution (>15 Pmolec/cm²) -46% (7.3 Pmolec/cm²). Total column (6 Pmolec/cm²): +4% (0.7 Pmolec/cm²) and -23% (5 Pmolec/cm²). V1.4.0 change: 33% and 60% reduction of tropospheric and total NO_2 biases.
	OFFL RPRO	NO_2 troposphere NO_2 stratosphere NO_2 total	-34% -5% -7%	2.7 Pmolec/cm² 0.3 Pmolec/cm² 1.5 Pmolec/cm²	
L2_HCHO	NRTI OFFL RPRO	HCHO, low HCHO, high	+27% -29%	8 Pmolec/cm² 15 Pmolec/cm²	HCHO bias varies with column amount: positive bias over clean areas (<2.5 Pmolec/cm²), negative bias over polluted areas (>8 Pmolec/cm²). No impact of version change to 2.1.4.
L2_SO2	NRTI	SO_2 column	0.2 DU	0.2 DU	Lack of validation sites in areas with high SO_2.
	OFFL	SO_2 column	0.2 DU	0.2 DU	
L2_CO	NRTI	CO column	6.5%	5%	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.
	OFFL	CO column	6.5%	5%	
L2_CH4	OFFL	CH4 column	-0.25%	0.69%	Along orbit stripes. Underestimation at low albedo. Remaining outliers with `qa_value` >0.5. 1-4% seasonal and SZA dependence of bias. Lower number of pixels with qa_value>0.5 since March 11 2020 due to changed cloud data.

Product ID	Stream	Product	Bias	Dispersion	Special features
L2_CLOUD	NRTI OFFL	CALv1 CTH (high) CALv2 CTH (high) CALv1 CTH (low) CALv2 CTH (low) CRBv1 CH (high) CRBv2 CH (high) CRB CH (low)	-30% -40% -15% -15% -20% -30% -40%	2 km 2 km 0.8 km 0.5 km 1.5 km 1.5km 0.5 km	Low clouds: CLOUDNET CTH<4km; high clouds: CLOUDNET CTH>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.
L2_AER_AI	NRTI	aerosol index	-1.1 AI unit	0.1 AI unit	Negative bias exceeding 1 AI unit after March 2019, attributed to irradiance data degradation.
	OFFL	aerosol index	-1.1 AI unit	0.1 AI unit	
L2_AER_LH	OFFL	aerosol layer height	50 hPa	100 hPa	Over ocean only. Larger bias and dispersion expected over land.

Table 2 – Representative quality indicators (bias, dispersion and special features) derived from the validation of the S5P TROPOMI data products listed in the **Table 1**, valid for all processor versions unless stated differently. CTH: cloud-top-height; CH: cloud height; COT: cloud optical thickness.

Table of Contents

1 Introduction

1.1 Background information on Sentinel-5 Precursor TROPOMI

TROPOspheric 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 km^2 before and 5.5 x 3.5 km^2 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 (O_3), nitrogen dioxide (NO_2), formaldehyde (HCHO), sulphur dioxide (SO_2), carbon monoxide (CO), methane (CH_4), 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_CoperESA]. As such, it has an operational and service-oriented vocation. With a 7-year 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 2023-2044.

1.2 Mission Performance Centre – Routine Operations Validation Service

Procured by an international consortium contracted by the European Space Agency (ESA), the S5P Mission Performance Centre (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 produces every three months the present *Routine Operations Consolidated Validation Report* (ROCVR). It also contributes quality assessment support to the evolution of 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 S5P MPC Routine Operations Validation Service. It integrates validation results from the MPC Validation Data Analysis Facility (VDAF) consortium (**Table 17**) with support from other activities and dedicated field campaigns documented on the TROPOMI website [ER_TROPOMI], as well as ad hoc contributions from S5P Validation Team (S5PVT) AO projects [ER_S5PVT].

Updated with a trimestral frequency, S5P data quality information provided in this document supersedes that provided in previous versions. It complements S5P data quality information provided in the *Product Readme Files* (PRFs) attached to S5P 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 S5P products and algorithms [ER_CoperATBD] and also on the TROPOMI Portal [ER_TROPOMI].

This update #11 of the ROCVR presents quality information for the S5P operational data products obtained in nominal mode from April 2018 until June 2021. For data products generated with the UPAS processor (O_3, HCHO, SO_2, clouds) it includes validation results for the processor upgrade to version 02.01.04 activated in Nov. / Dec. 2020. For the NO_2 data products generated with the NLL2DP processor it includes initial validation results for the processor upgrade to version 01.04.00 activated on November 29 (OFFL) and December 2 (NRTI), 2020. This document is structured as follows:

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.

Product ID	Level-2 Geophysical Quantity	Requirement: Vertical Resolution	Requirement: Bias	Requirement: Uncertainty
L2_O3	Total O_3	total column	5%	2.5%
L2_O3_PR	O_3 profile (incl. troposphere)	6 km	30%	10%
L2_O3_TCL	O_3 tropospheric column	tropospheric column	25%	25%
L2_NO2	NO_2 tropospheric column	tropospheric column	50%	0.7 Pmolec.cm^{-2}
	NO_2 stratospheric column	stratospheric column	10%	0.5 Pmolec.cm^{-2}
L2_SO2	Total SO_2	total column	0.5 DU	1 DU
	Enhanced SO_2 (SCD <1.5 DU)	total column	0.5 DU	1 DU
	Enhanced SO_2 (SCD >1.5 DU)	total column	30%	30%
L2_HCHO	Total HCHO	total column	80%	12 Pmolec.cm^{-2}
L2_CO	Total CO	total column	15%	10%
L2_CH4	Total CH_4	total column	1.5%	1%
L2_CLOUD	Cloud Fraction	total column	20%	0.05
	Cloud Height (pressure)	total column	20%	0.5km (P<30hPa)
	Cloud albedo (optical thickness)	total column	20%	0.05 (10)
L2_AER_AI	Aerosol Absorbing Index	total column	1 AAI	0.1 AAI
L2_AER_ALH	Aerosol Layer Height	total column	100 hPa	50 hPa

Table 3 – Data quality targets for the operational Sentinel-5 Precursor TROPOMI Level 2 data products: measurement bias and (random) measurement uncertainty (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 product identified in **Table 4**.

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

Product	Stream	Version	In operation from	In operation until
L1B_RA1/…/8		01.00.00	orbit 2818, 2018-04-30	current version
L1B_IR_UVN/SIR		01.00.00	orbit 2818, 2018-04-30	current version

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

3.2 Recommendations for data usage followed

The product is stored as NetCDF4 file. The NetCDF4 file contains both the data and the metadata for the product.

For OFFL and RPRO data the product is stored as a single file per satellite orbit, for NRTI data the product is stored as multiple files per orbit.

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. Details of the data format are provided in the Input/Output Data Specification. The metadata contained in the L1b data products are described in the Metadata Specification. 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 have been applied:

- 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 8 bands are processed individually in the L0-1b data processor. In case of missing data, for example in case of data drop-outs 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 S5P TROPOMI L1B data products are also compared to modelling output and to other satellite measurements, specifically from EOS-Aura OMI and from Suomi-NPP OMPS.

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 products shows agreement of 0.02 to 0.04 nm, which is within the pre-launch calibration accuracy.

Initial validation of the L1B_RA reflectance with respect to OMI and OMPS data indicates that TROPOMI is within 5% for the shorter wavelengths in band 3 and improving to 2% towards the longer wavelengths in band 4. For the short wave UV in band 1 TROPOMI is within 8% +/-2% of the expected modeled reflectance.

In general radiometric errors in bands 1 and 3 are large but they vary slowly over wavelength and most L2 retrievals are insensitive to such errors. Additional validation indicates that for bands 3 to 7 the mission requirements for the reflectance are met if the uncertainty of the method of 3 to 5% is taken into account.

The largest source of error in the reflectance is due to the initial pre-launch irradiance calibration. This is a known issue and will be addressed in future updates.

The validation of the TROPOMI L1B_IR irradiance product shows that it is within 3 to 10% depending on the used reference spectrum and that there is a radiometric mismatch between band 2 and 3.

Additional validation with other solar irradiance spectra concludes that the difference shows a smooth wavelength dependence, most likely caused by optical setup effects during the on-ground calibration. This anomaly affects the UV and UVIS channels, and can be corrected for in the update of the L01b processor. After this correction the differences with reference spectra reduces to 2% and is within the expected radiometric accuracy. For the NIR and SWIR channels the difference shows no wavelength dependence but an offset that is within the radiometric accuracy budget. For these channels no correction is foreseen.

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).

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 S5P Validation Team (S5PVT) 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 *S5P First Public Release Validation Workshop* (ESA/ESRIN, June 25-26, 2018), at the *3rd S5PVT workshop* (ESA/ESRIN, November 11-14, 2019), and at the *4th S5PVT 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-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. 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 02.01.03 in July 2020.

Current conclusions are valid for the S5P data obtained in the operational phase E2 of the mission, from May 2018 until May 2021, and on the reference data available at the time of this report: typically, until the end of April 2021 for the Dobson and Brewer data, and up to the end of May 2021 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 200 to 1600 co-locations have been identified at about 40 Brewer and Dobson sites and at 12 ZSL-DOAS SAOZ sites, sampling many latitudes from the Arctic to the Antarctic (**Figure 1**).

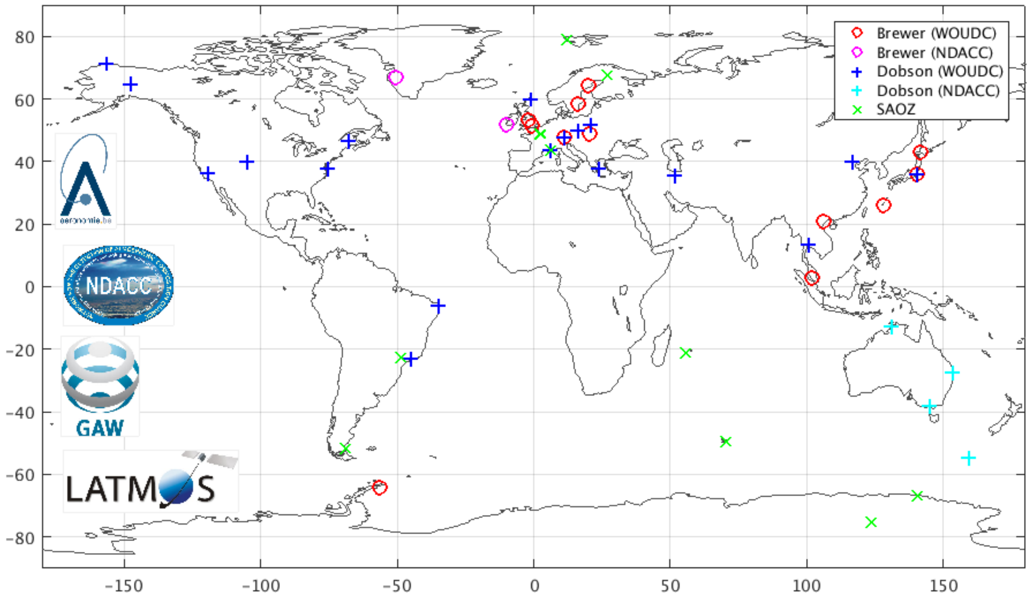


Figure 1: Geographical distribution of Brewer, Dobson and ZSL-DOAS ground-based stations for which co-locations with S5P L2_O3 NRTI ozone data have been used (May 2018 until February 2021).

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%).

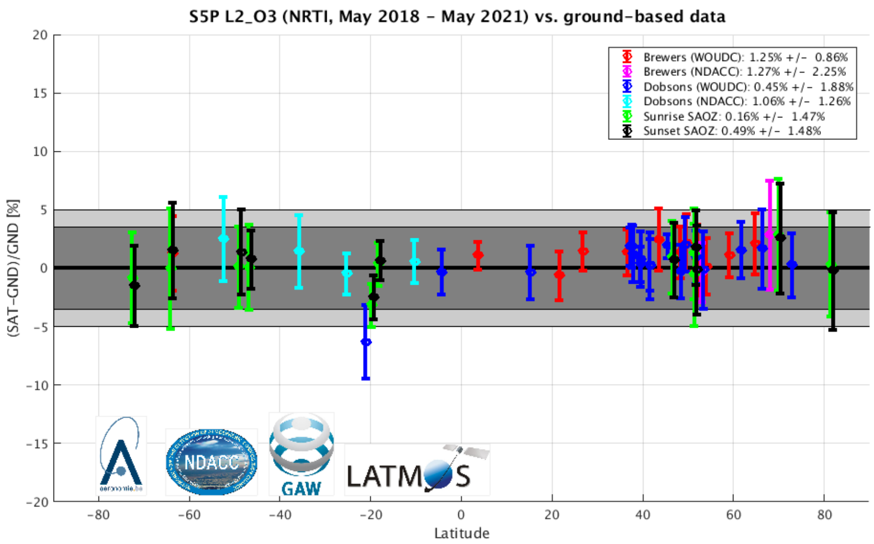


Figure 2: Meridian dependence of the median (the circular markers) and spread (±1 sigma, the error bars) of the percent relative difference between S5P TROPOMI L2_O3 (PDGS NRTI processor v1.0.0 up to v2.1.4, May 2021) and ground-based (GND) ozone column data, represented at individual stations from the 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).

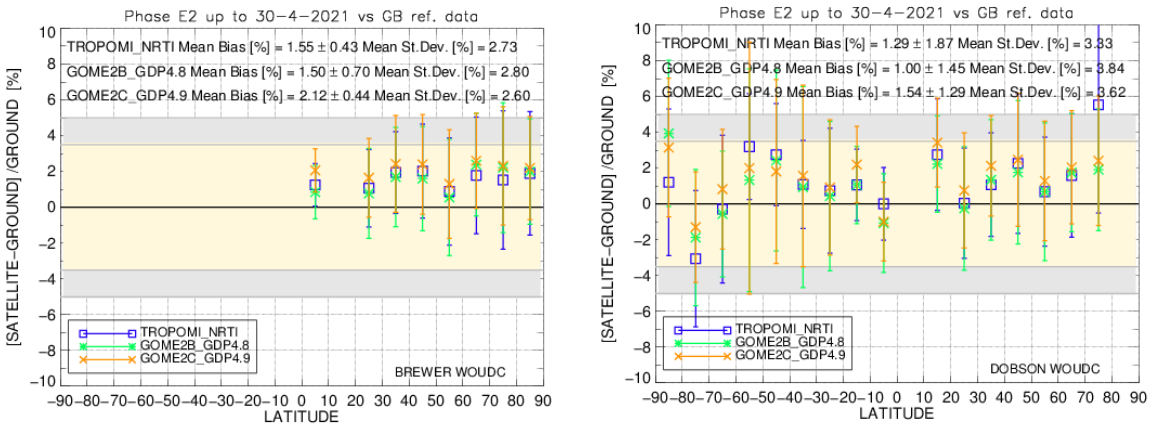


Figure 3: Comparison of the mean percentage differences between three satellite products (S5P TROPOMI L2_O3 NRTI, GOME-2B GDP 4.8, and GOME-2C GDP4.9) and ground-based total ozone data, versus latitude. The Brewer network comparisons are shown in the left-hand panel and the Dobson network comparisons in the right-hand panel. Both ground-based datasets are downloaded from the WOUDC. Time period: is May 2018 – April 2021.

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

On 6 August 2019, the nominal ground pixel resolution of the TROPOMI measurement was reduced to 5.5 x 3.5 km^2, 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). **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 Sodankyla (at the top in **Figure 4**), sites 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.

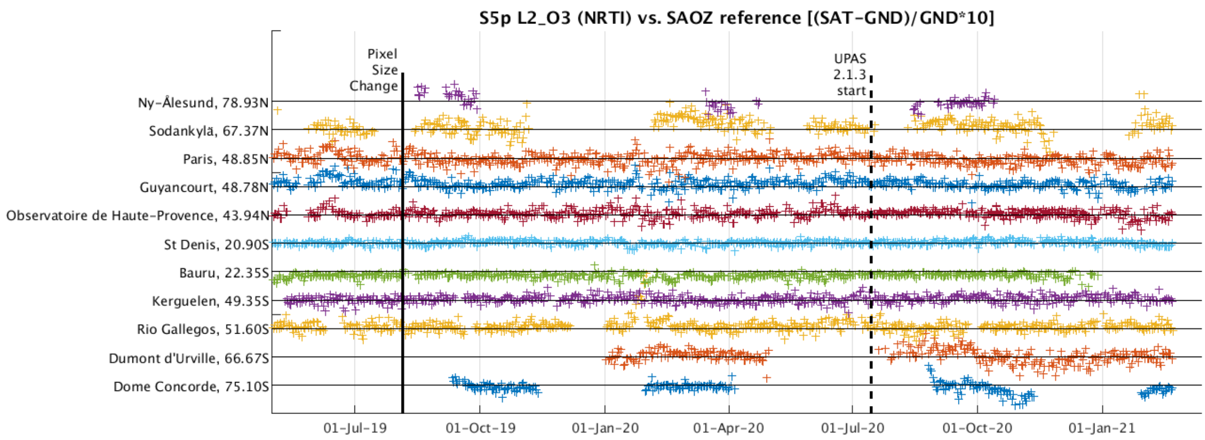


Figure 4: Time series of the difference between the S5P NRTI L2_O3 ozone column data and the SAOZ correlative data at 11 stations from pole to pole, from early 2019 to May 2021. The moment of the switch in TROPOMI pixel size (along track reduction from 7 to 5.5 km) is identified by the bold vertical line, the UPAS processor upgrade to v02.01.03 by the dashed line.

4.3.4 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 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 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 χ_r^2 , defined as:

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

where, σ_{SAT} and σ_{GND} refer to the ex-ante satellite and ground-based measurements respectively. The histogram of these χ_r^2 is shown in **Figure 5**. The histogram peaks close to one, with only one value near 3. This suggests the ex-ante uncertainties to be realistic.

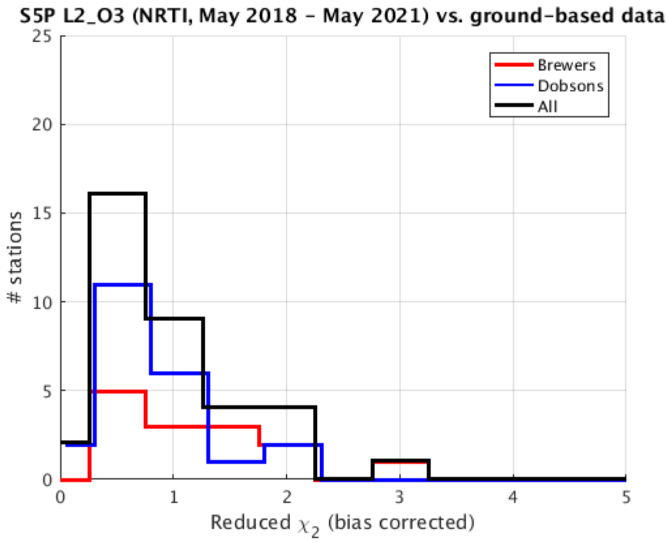


Figure 5: Histogram of the χ_r^2 values (one per reference instrument) for S5p NRTI total ozone versus different ground-based instruments.

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 6**).

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.

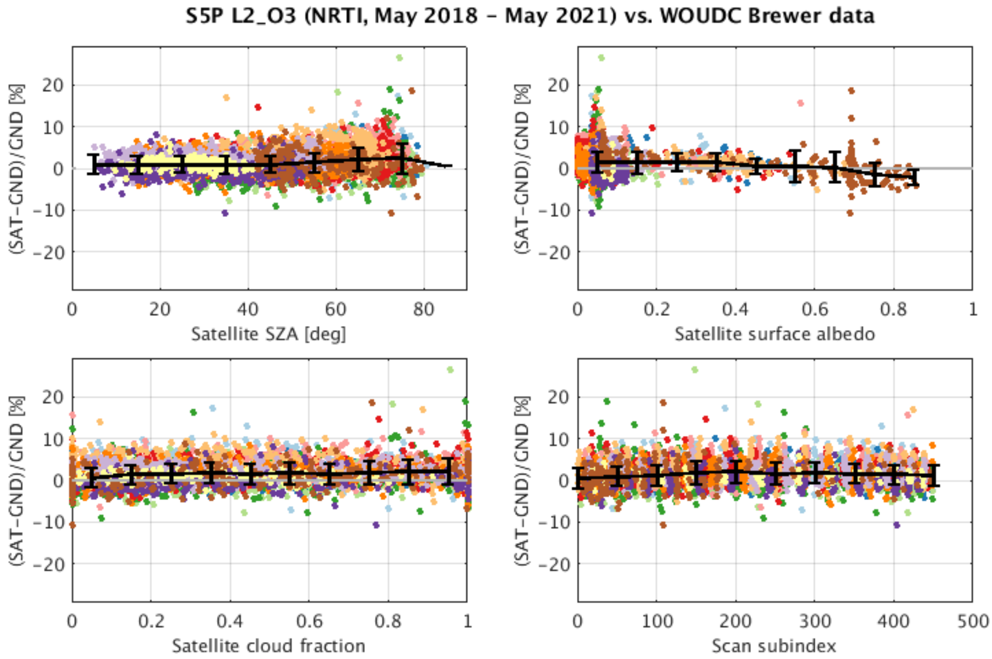


Figure 6: Dependence of the difference between S5P NRTI and ground-based Brewer total ozone data on the solar zenith angle (SZA), the surface albedo, the fractional cloud cover, and the scan sub-index of the satellite measurement. Black curve: mean and standard deviation over bins of 10 degrees of SZA, 0.1 of surface albedo, 0.05 of cloud fraction, and 50 pixels of the across-track scan.

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 7**. The overall good agreement is corroborated by Pearson correlation coefficients always above 0.95.

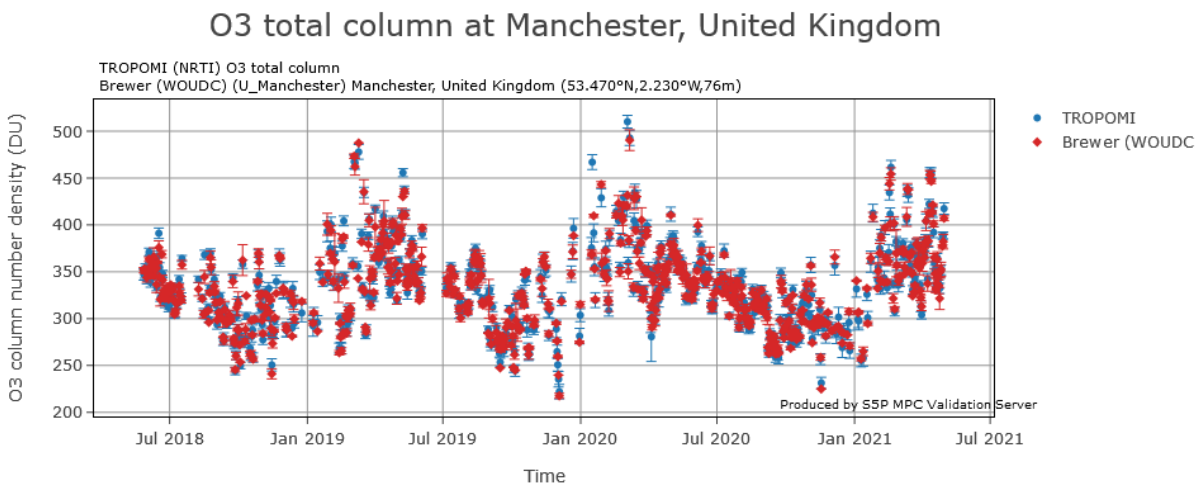


Figure 7: Time series of S5P TROPOMI NRTI and Brewer total ozone data at the station of Manchester in the United Kingdom (data courtesy J. Rimmer, University of Manchester).

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 8** where NRTI and OFFL data are compared.

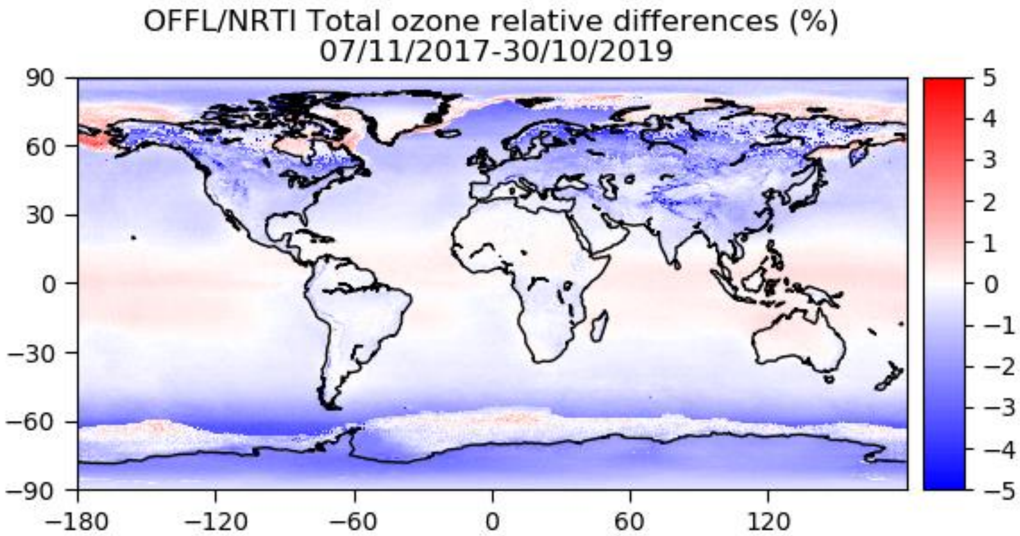


Figure 8: Percent relative difference between S5PL2_O3 total ozone data retrieved with the NRTI and OFFL processors (period November 2017 through October 2019). Geographical patterns in this case cannot be associated with real ozone features but reveal rather the effect of using either a surface albedo 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).

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_columnn 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 S5P Validation Team (S5PVT) 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 *S5P First Public Release Validation Workshop* (ESA/ESRIN, June 25-26, 2018), at the *3rd S5PVT Workshop* (ESA/ESRIN, November 11-14, 2019), and at the 4th S5PVT 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-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 valid for the S5P data obtained in the operational phase E2 of the mission, from May 2018 until May 2021, and on the reference data available at the time of this report: typically, until April 2021 for the Dobson and Brewer data, and up to the end of May 2021 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 200 to 1600 co-locations have been identified at about 40 Brewer and Dobson sites and at 12 ZSL-DOAS SAOZ sites, sampling many latitudes from the Arctic to the Antarctic (**Figure 9**).

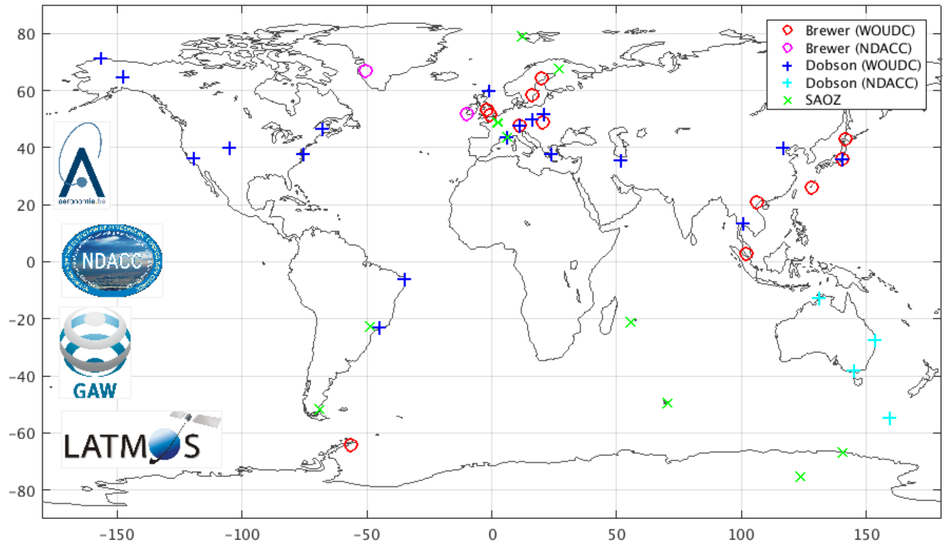


Figure 9: Geographical distribution of Brewer, Dobson and ZSL-DOAS ground-based stations for which co-locations with S5P L2_O3 OFFL ozone data have were used (period up to May 2021).

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 10**. 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 11**). This median bias value falls well within the mission requirements (max. bias 5%).

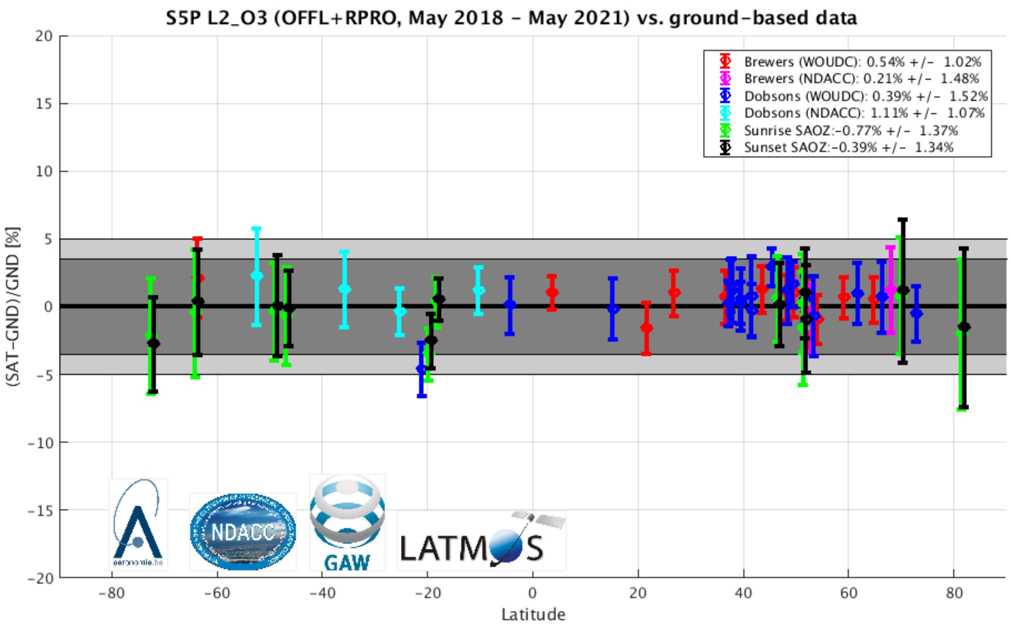


Figure 10: Meridian dependence of the median (the circular markers) and spread (±1 sigma, the error bars) of the percent relative difference between S5P TROPOMI L2_O3 (PDGS OFFL processors v1.1.7, v1.1.8, v2.1.3, and v2.1.4) and ground-based (GND) ozone column data, represented at individual stations 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).

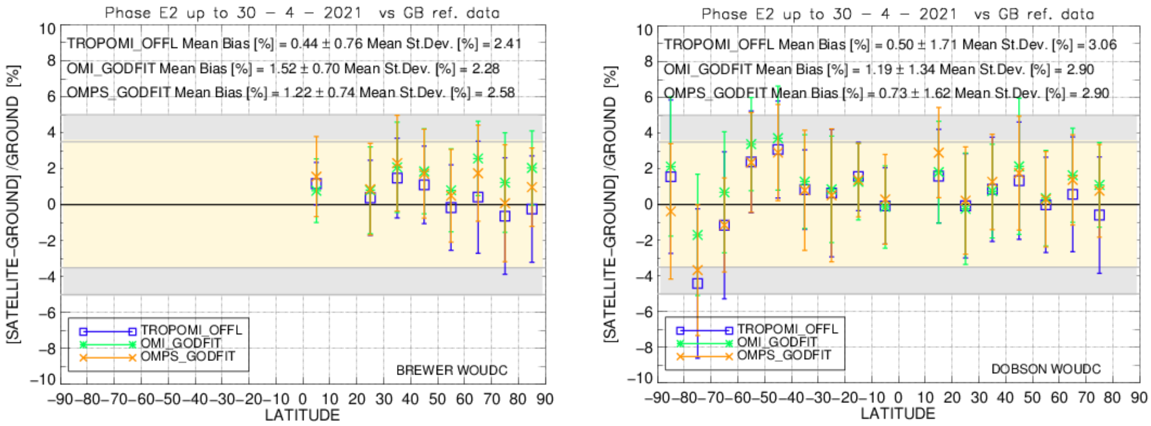


Figure 11: Comparison of the mean percentage differences between three satellite data products (S5P L2_O3 OFFL, OMI GODFIT v4 and OMPS GODFIT v4) and ground-based total ozone data, versus latitude. The Brewer network comparisons are shown in the left-hand panel and the Dobson network comparisons in the right-hand panel. Both datasets are downloaded from the WOUDC. The time period of data used for these plots is May 2018 – April 2021.

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

On 6 August 2019, the nominal ground pixel resolution of the TROPOMI measurement was reduced to 5.5 x 3.5 km^2, 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). **Figure 12** 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 new processor upgrade will require a more extensive set of co-locations and is planned for a future quarterly update of this report.

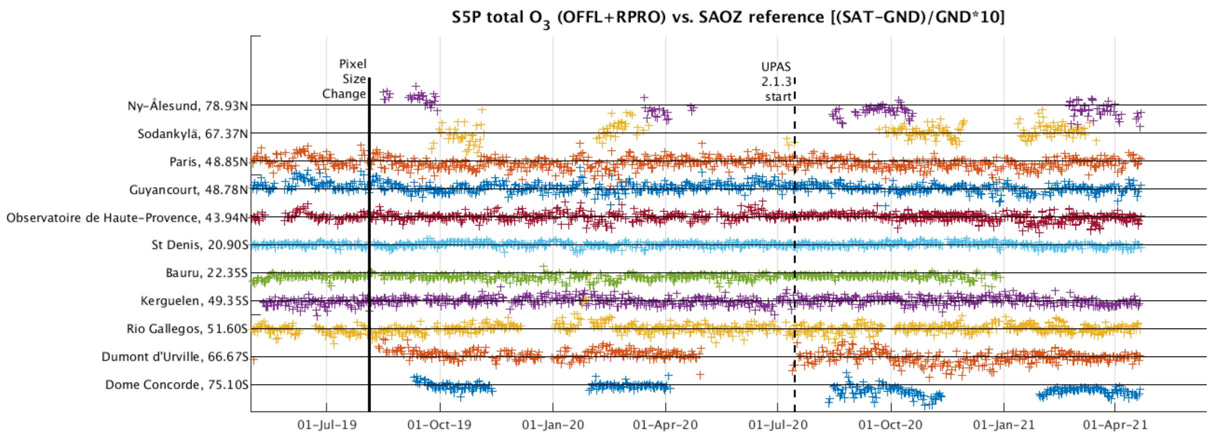


Figure 12: Time series of the difference between the S5P OFFL L2_O3 ozone column data and the SAOZ correlative data at 11 stations from pole to pole, from early 2019 to May 2021. The moment of the switch in TROPOMI pixel size (along track reduction from 7 to 5.5 km) is identified by the bold vertical line, the UPAS processor upgrade to v02.01.03 by the dashed line.

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 10**). 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 χ_r^2 , defined as:

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

where, σ_{SAT} and σ_{GND} refer to the ex-ante satellite and ground-based measurements respectively. The histogram of these χ_r^2 is shown in **Figure 13**. The histogram peaks close to one, with only a few values above 3. This suggests the ex-ante uncertainties to be realistic.

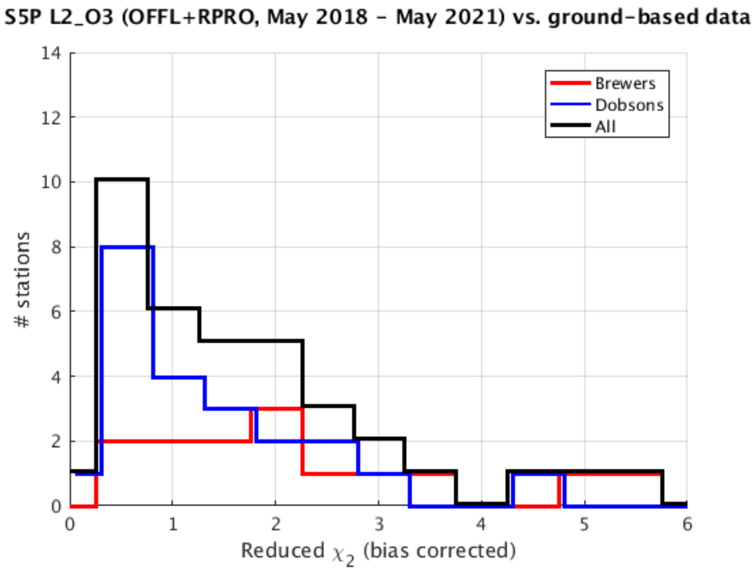


Figure 13: Histogram of the χ_r^2 values (one per reference instrument) for S5p OFFL total ozone versus different ground-based instruments.

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 14**).

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. Moreover, satellite-to-satellite comparisons indicate some systematic differences at the largest SZAs, up to 5%, as illustrated in **Figure 15**.

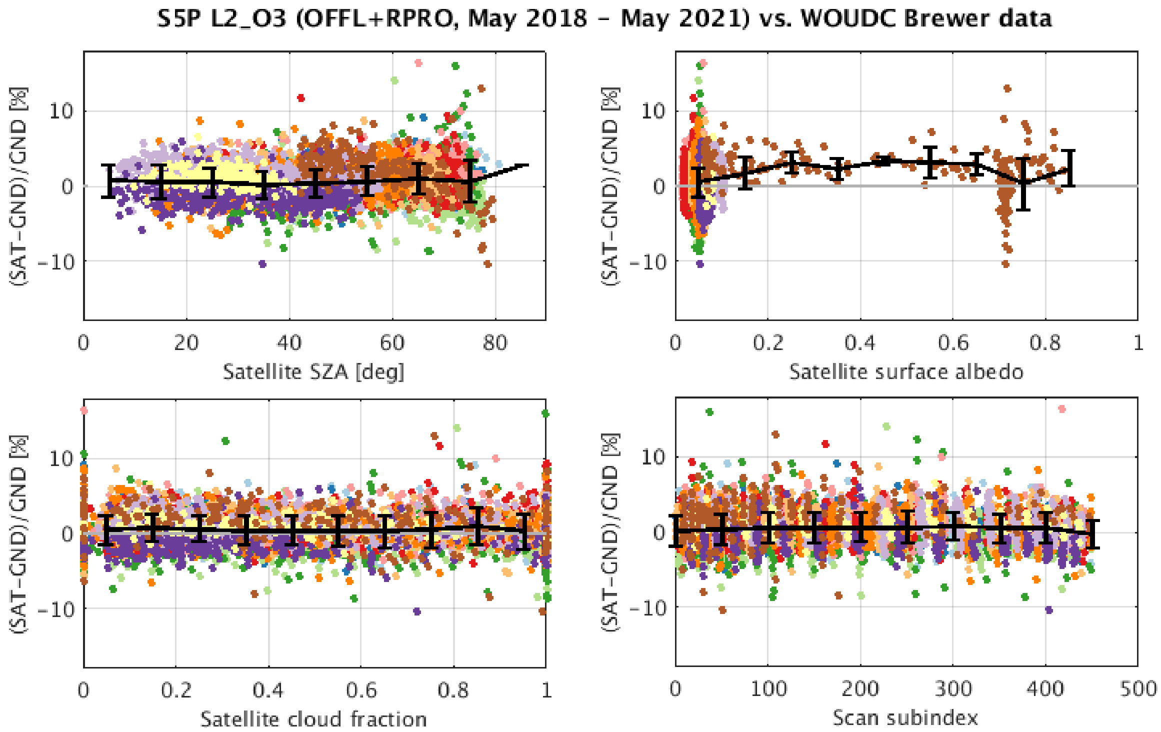


Figure 14: Dependence of the difference between S5P OFFL and ground-based Brewer total ozone data on the solar zenith angle (SZA), the surface albedo, the fractional cloud cover, and the scan sub-index of the satellite measurement. Black curve: mean and standard deviation over bins of 10 degrees of SZA, 0.1 of surface albedo, 0.05 of cloud fraction, and 50 pixels over across-track scan.

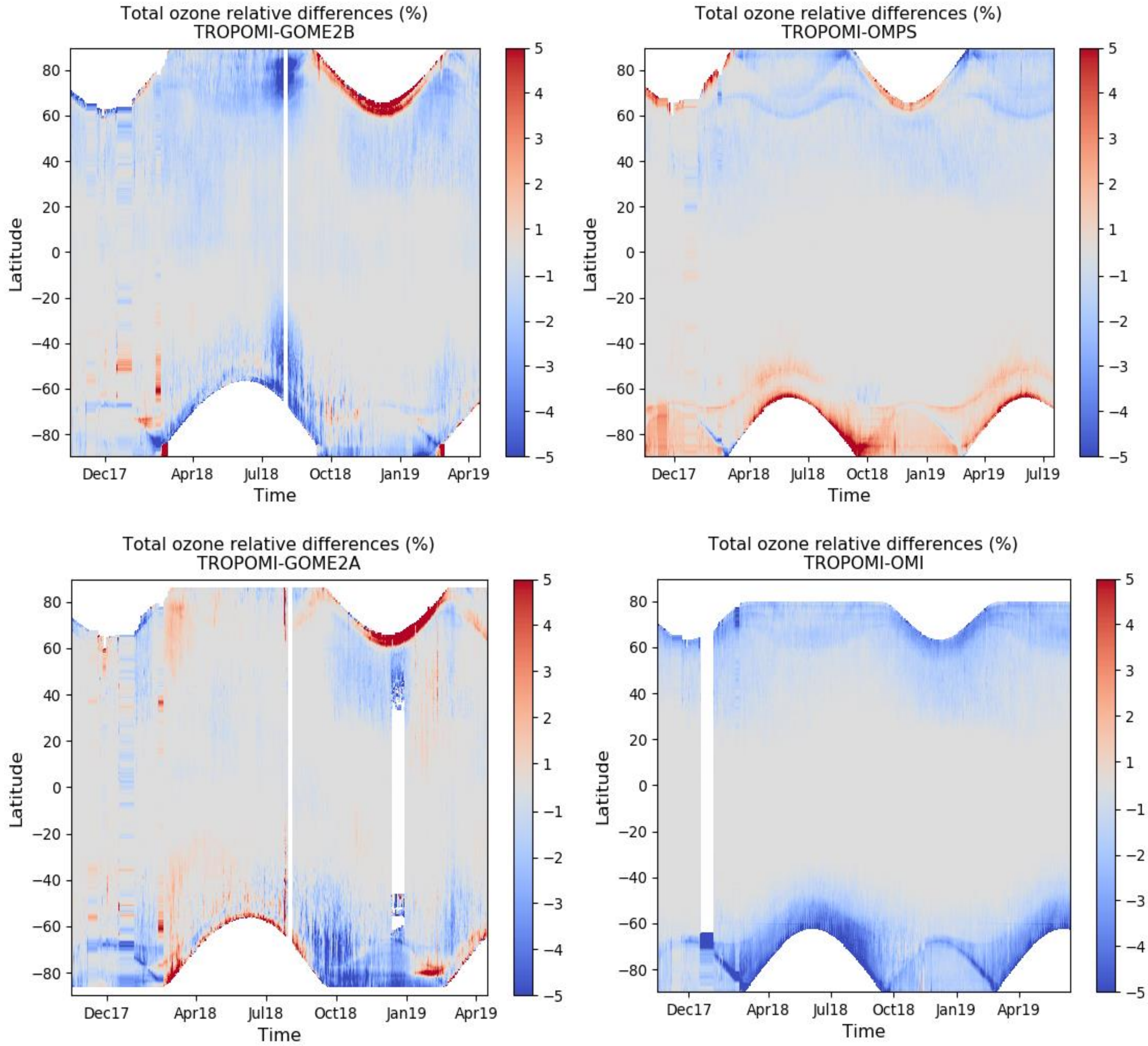


Figure 15: Latitude and time cross-section of the percent relative difference between S5P L2_O3 OFFL total ozone data and comparable total ozone data sets from other satellite: GOME-2B GDP 4.8, S-NPP OMPS, GOME-2A GDP 4.8, and Aura OMI.

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 16**. The overall good agreement is corroborated by Pearson correlation coefficients always above 0.95.

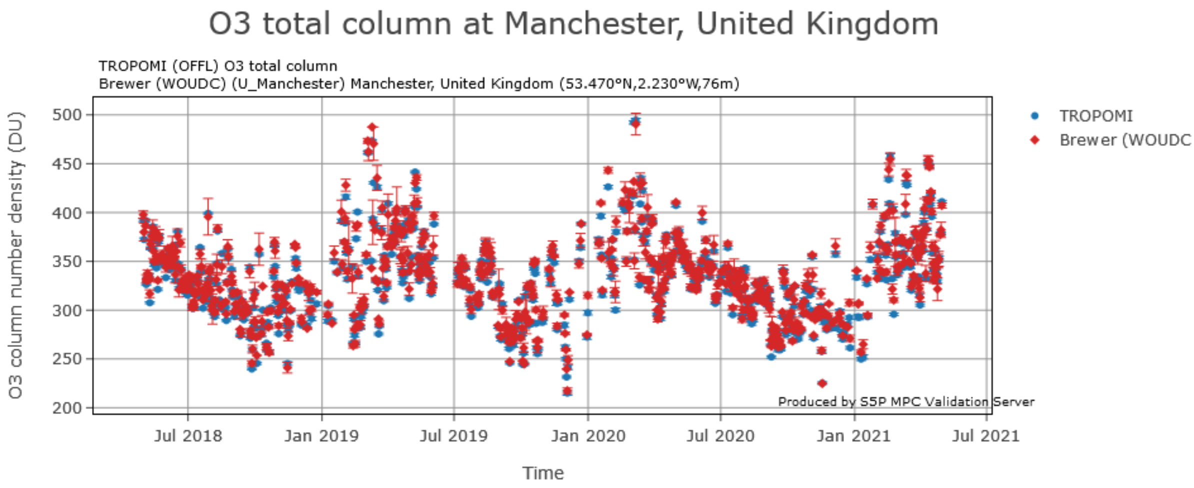


Figure 16: Time series of S5P TROPOMI OFFL and Brewer total ozone data at the station of Manchester in the United Kingdom (data courtesy J. Rimmer, University of Manchester).

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 8** where NRTI and OFFL data are compared. See Sect. 4.3.7 for more details. Moreover, **Figure 15** highlights patterns associated with differences in the dependence on SZA already discussed in Section 4.4.5.

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. They represent a three-day average 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 sites of the ground-based SHADOZ network (ER_SHADOZ) are compared routinely to S5P data (see Hubert *et al.*, 2020). The SHADOZ data version used here is V06. The ozonesonde profile data are first quality controlled (Hubert *et al.*, 2016, 2020) 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 19** and **Figure 21**). 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°x2.5° to 1°x2°.

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 S5P 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 S5P Validation Team (S5PVT) AO projects. This summary is based on coordinated operational validation activities carried out using the Automated Validation Server of the S5P 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 – 18 May 2021, the ground-based validation analysis considers 1102 S5P OFFL CCD data products and 752 ozonesonde flights at nine sites across the tropics (**Figure 17**). S5P data averaged over the entire tropical region are also compared (**Figure 20**) to GOME-2B data (May 2018 – November 2020) and to OMI data (May 2018 – December 2020).

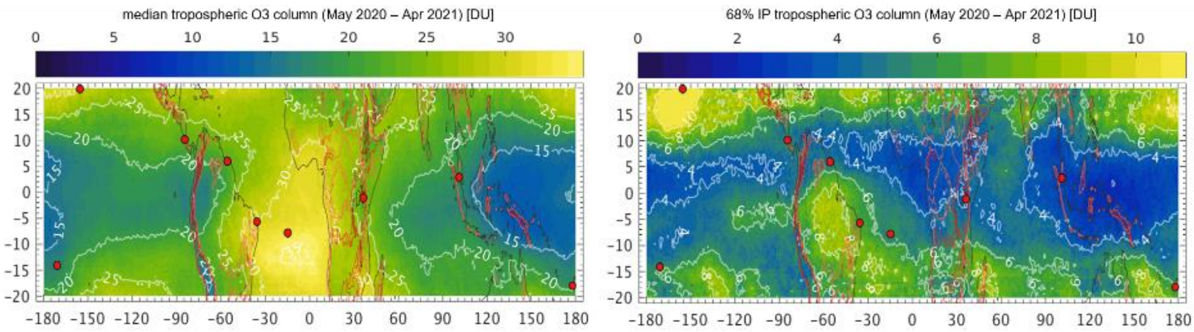


Figure 17: Median value (left) and half-width of 68% interpercentile (right) of S5P OFFL tropospheric ozone column data (CCD) over a full year of operations (May 2020 – April 2021). Red markers locate the nine ground-based ozonesonde stations used in the validation analysis. Red contours indicate surface elevation (500, 1000, 2000 m). These maps provide context to **Figure 18** and **Figure 19**.

5.3.3 Bias

S5P tropospheric O_3 column values are on average larger than the ozonesonde values at all nine sites (**Figure 19** and **Figure 21**). The mean bias over the network is +16% or +3.1 DU (**Figure 21**, 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 20**. The agreement with OMI is good, with a mean difference of +0.2 DU or +0.5%. 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.

5.3.4 Dispersion

The half 68% interpercentile of the difference (between S5P and ozonesonde data) ranges within 16-36% or 3.5-7.9 DU (**Figure 19** and **Figure 21**), and the network average is 25% or 4.8 DU (**Figure 21**, centre and bottom right). Dispersion values at four sites are not compliant with the mission requirement for the random component of the uncertainty (<25%). However, three of these sites are located in an area with large natural percentage variability in the tropospheric O_3 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 20** (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 20**). 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.*, 2020). 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 processor version early July.

S5P time series in **Figure 18** show signatures of biomass burning affected air (high tropospheric O_3 values) over Atlantic and African sites (Heredia, Paramaribo, Natal, Ascension Island, Nairobi), and of intense convective activity (low tropospheric O_3 values) over Pacific stations (Samoa, Suva, Sepang Airport). Comparisons during the biomass season show systematically more elevated positive biases over Paramaribo and Heredia. The temporary, additional bias amounts to, on an average, about 25% or 5 DU. The pattern is particularly clear at Paramaribo, occurring every year since 2018. The recently released 2020 Heredia ozonesonde data confirm what was noted in 2018, while the pattern was not as convincing for the 2019 season at this site. At the Atlantic sites in the Southern Hemisphere (Natal, Ascension Island, Nairobi), however, there is no consistent, annual repetition of more elevated biases during the biomass burning season. We conclude that, at this point, there is no unequivocal evidence for a bias pattern during biomass burning conditions across the entire tropical belt. However, it cannot be excluded either that such a bias pattern is a local effect. This issue will be monitored by the validation team.

Co-located S5P and reference measurements correlate fairly well for sites with well-sampled comparison time series. Pearson's correlation coefficients range between 38% and 75% at individual stations, while the network average is 61% (**Figure 21**, top left).

5.3.7 Geographical patterns

Annual median TROPOMI data (May 2020 – April 2021, **Figure 17**) 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 O_3 data record as well. An in-depth analysis can be found in Hubert *et al.* (2020).

5.3.8 Preliminary validation of UPAS processor upgrade to version 02.01.03

The upgrade to the new Level-2 processor in the ground segment occurred on 16th July 2020. L2_O3_TCL data products starting from 5th July 2020 are generated using the new processor. By the time of this report 313 S5P products were processed using 02.01.03 and 02.01.04, leading to 163 co-locations with ozonesondes at six sites (Hilo, Heredia, Paramaribo, Ascension Island, Samoa, Suva). No clear change was seen in the comparisons (**Figure 19**) although confidence remains limited since less stations provide data in the latter period and there is considerable dispersion in the comparisons.

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 18**) 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 S5P ground pixel size (on 6 August 2019) does not have a noticeable impact on the quality of S5P tropospheric ozone products (**Figure 18** and **Figure 19**). Estimates of bias and dispersion before and after the change are consistent.

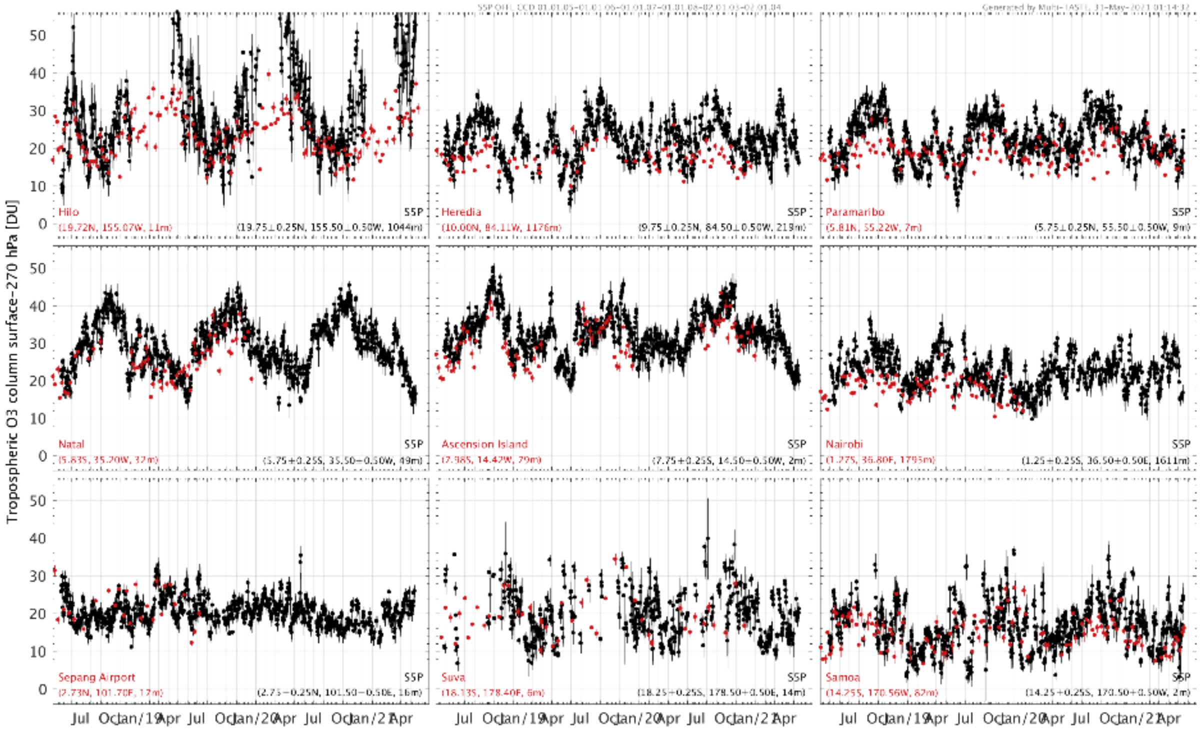


Figure 18: Time series of spatially co-located tropospheric O_3 column data by ozonesonde (red) and by S5P OFFL v01.01.05-v02.01.04 (black). All data were screened following recommendations by the data providers.

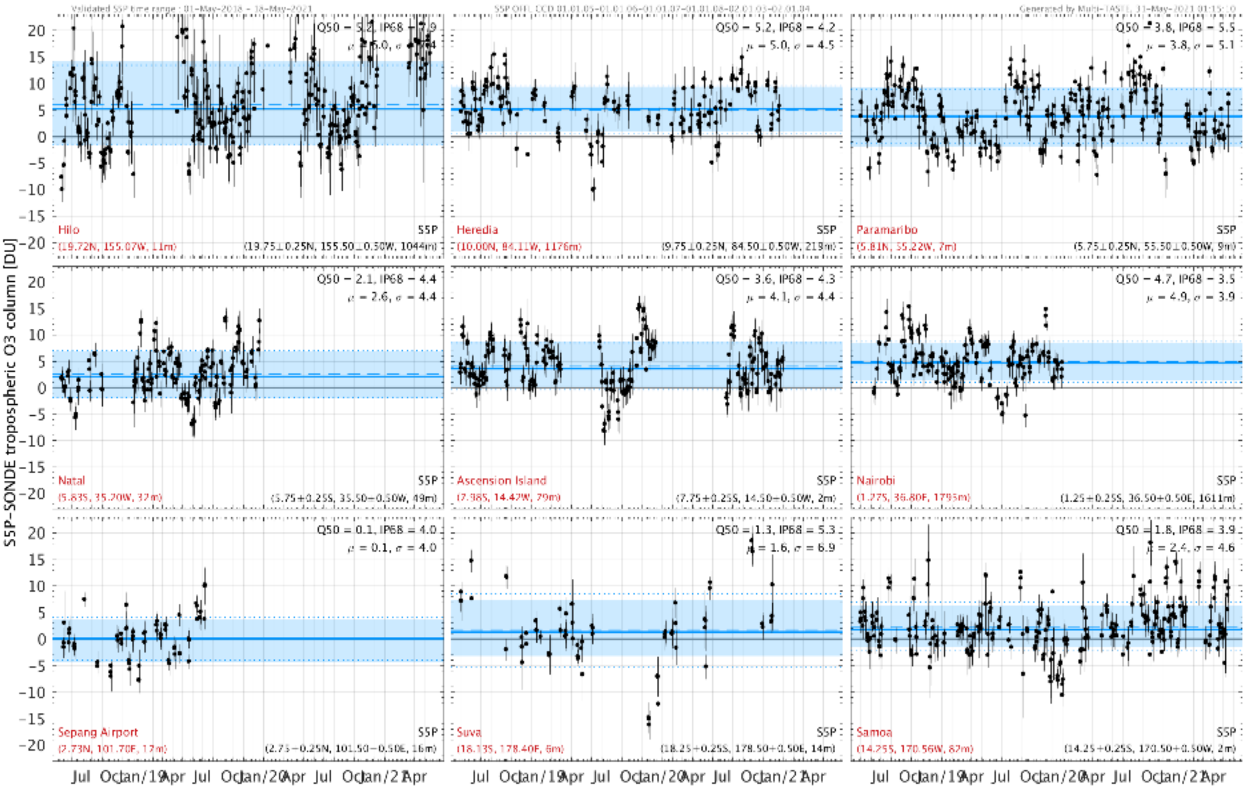


Figure 19: Time series of the absolute difference between spatially and temporally co-located S5P and ozonesonde tropospheric O_3 column data. The blue line and shaded area shows the median value and the range between the 16% and 84% percentiles. Positive values indicate a high bias of S5P w.r.t. the reference.

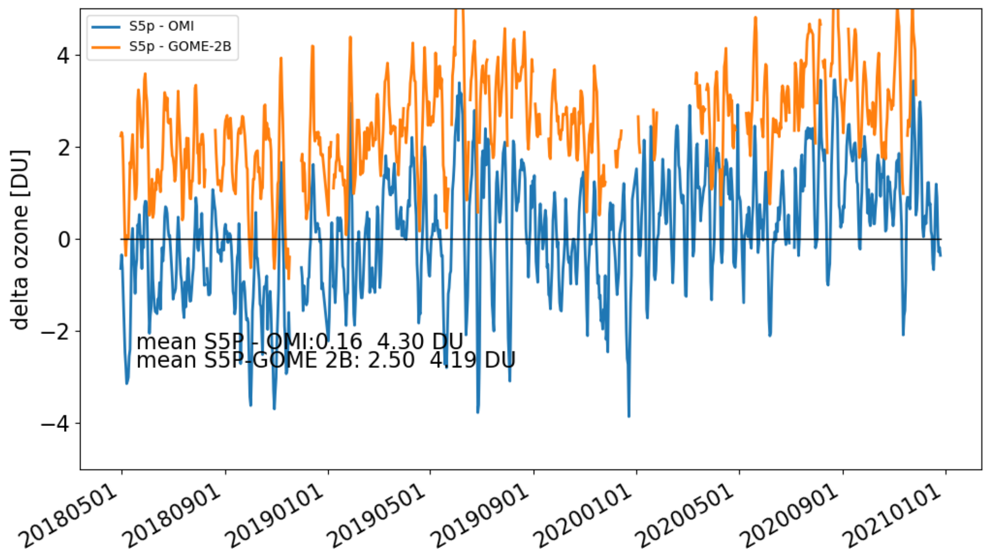


Figure 20: Difference time series of daily tropospheric O_3 column data averaged over the 20°S – 20°N tropical belt. S5P OFFL CCD data are compared to satellite data by OMI and GOME-2B, positive values indicate a high bias of S5P w.r.t. the reference.

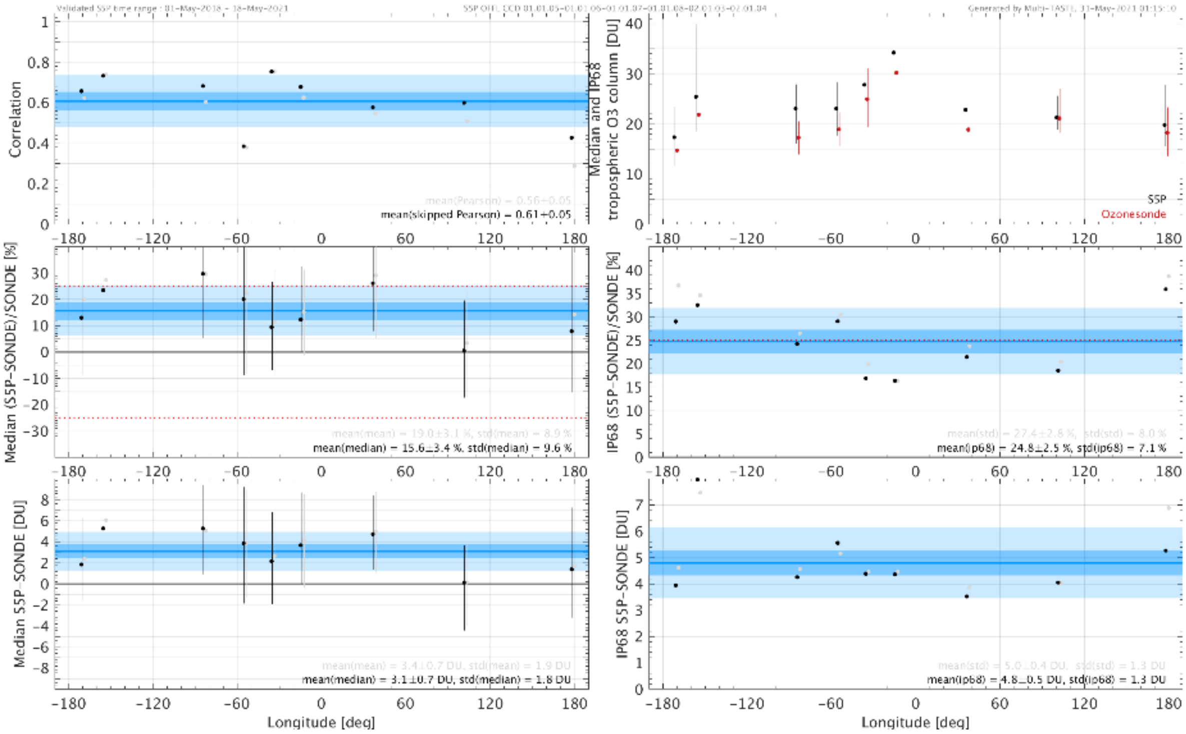


Figure 21: Overview of correlation (top left), median bias (middle & bottom left) and dispersion of the difference (middle & bottom right) of S5P tropospheric O_3 column data for each SHADOZ site (black markers). Black vertical bars represent the 68% interpercentile of the difference. The mean, standard 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.

6 Validation Results: L2_NO2

6.1 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 30 April 2018. The NRTI data covers the full range of versions from 01.00.01 to 01.04.00 as there is no reprocessing necessary. 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 six times from 01.02.00 up to the current version 01.04.00 that started in 2020/11/29.
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 6.4 demonstrates evidence that NRTI and OFFL data do not differ significantly and that their respective validations yield similar conclusions.

6.2 Validation approach

6.2.1 Ground-based monitoring networks

Tropospheric NO2 – 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 bias of maximum 20% and a precision better than 30% at this set of stations. The validation with NIDFORVAL MAX-DOAS data 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 EVDC (through conversion to fully GEOMS and HARP compatible data) lead to the inclusion of new sites to the VDAF-AVS and several time-period updates. Careful revision (and eventual reprocessing) are ongoing for the Thessaloniki sites and the UNAM sites (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 22**, 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.

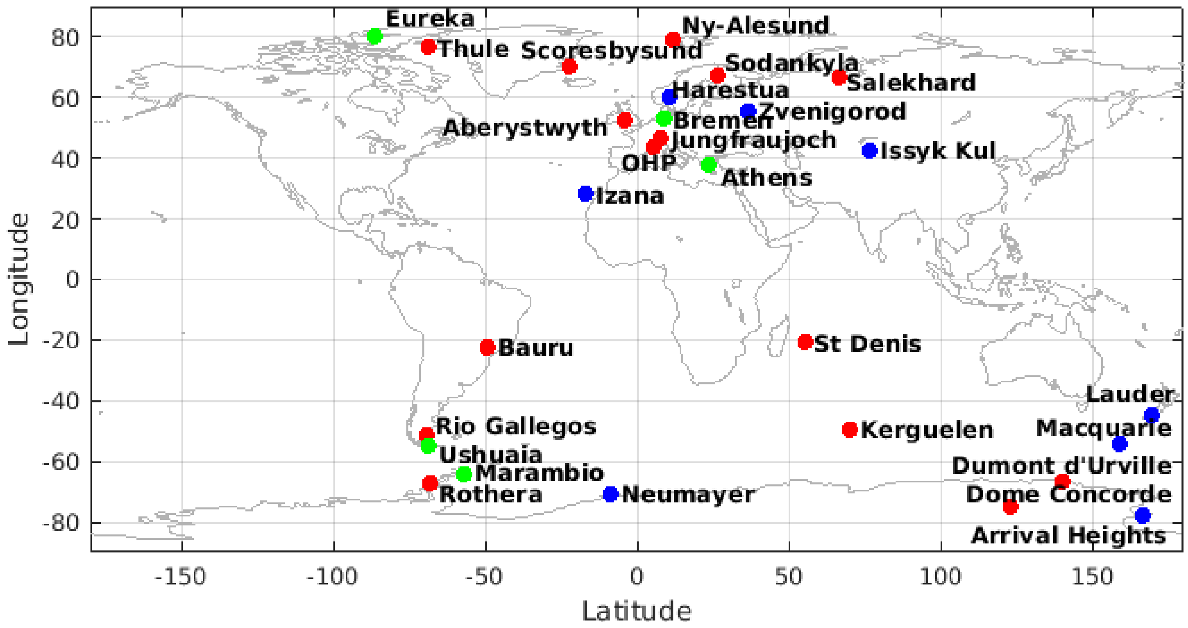


Figure 22: Geographical distribution of the NDACC ZSL-DOAS instruments measuring routinely stratospheric NO_2 and yielding co-locations with the current S5P L2_NO2 datasets. Stations marked with a red dot contribute fast delivery data coming from the LATMOS_RT facility. Blue and green dots depict the NDACC stations that contributes ZSL-DOAS data directly through the NDACC DHF and the AO project NIDFORVAL, respectively.

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. 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 on the order of 10%. To reduce mismatch errors due to the significant difference in horizontal smoothing between S5P and ZSL-DOAS measurements, S5P NO_2 values (from ground pixels at high resolution) are averaged over the footprint of the air mass to which the ground-based zenith-sky measurements are sensitive. Some 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).

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 NO2 data have a bias of maximum 10-15% and a precision of roughly 0.28 Pmolec/cm² (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 closest Pandora measurement closest in time is selected, with a maximum time difference of 30 min. If the Pandora instrument operates at an elevated site 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.

6.2.2 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 onboard the EOS-Aura satellite that was launched in July 2004.

6.2.3 Field campaigns and modelling support

None for this report.

6.3 Validation of L2_NO2

6.3.1 Recommendations for data usage

In order to avoid misinterpretation of the data quality, it is recommended to only use TROPOMI NO_2 measurements with a `qa_value` ≥ 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` ≥ 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.

6.3.2 Status of validation

This section presents a summary of the key validation results obtained by the MPC VDAF and by S5PVT AO projects. 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-UB, and the HARP toolset (version 1.6).
An in-depth discussion of the routine validation results up to March 2020 is now published (Verhoelst 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. In summary, the FRM data streams cover the time period from May 2018 to May 2021:

- MAX-DOAS tropospheric columns from the NIDFORVAL AO project from 25 stations. Vertical smoothing harmonization could be applied for 12 stations. NDACC MAX-DOAS tropospheric column data at the VDAF Automated Validation Server via EVDC are available from 7 stations (De Bilt, Cabauw (KNMI), Uccle, Xianghe (BIRA-IASB), Bremen, Athens (IUP-B), and Mainz (MPIC)).
- NDACC ZSL-DOAS stratospheric columns from 25 stations sample latitudes between 80°N (Eureka) and -75°S (Dome C).

- PGN total column direct-sun data are available from 35 sites at the VDAF Automated Validation Server sample the latitude range from 78°N to -46°S.

6.3.3 Tropospheric NO_2 column

6.3.3.1 Bias

The OFFL NO_2 tropospheric column values are compared with data from 25 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 23** and **Figure 24**. The median bias for the subset of 7 VDAF-AVS MAX-DOAS stations (inspection 2021/05/25, 3005 co-locations) is -1.35 Pmolec/cm² (-26%). These results are within the mission requirement of a maximum bias of 50%.

With a site-to-site dispersion (NIDFORVAL, IP68/2 over all site medians) of 2.6 Pmolec/cm², a single bias number for all sites has limited meaning. Three regimes can be identified for comparisons, as discussed in Verhoelst (2021): (1) low tropospheric NO_2 values (below 2 Pmolec/cm²), (2) polluted sites (from 3 to 14 Pmolec/cm²), and (3) extremely polluted sites (median tropospheric column larger than 15 Pmolec/cm²). The median bias in these regimes is about 0.1 Pmolec/cm² (10.5%), -1.9 Pmolec/cm² (-34.3%), and -10 Pmolec/cm² (-45.6%), respectively.

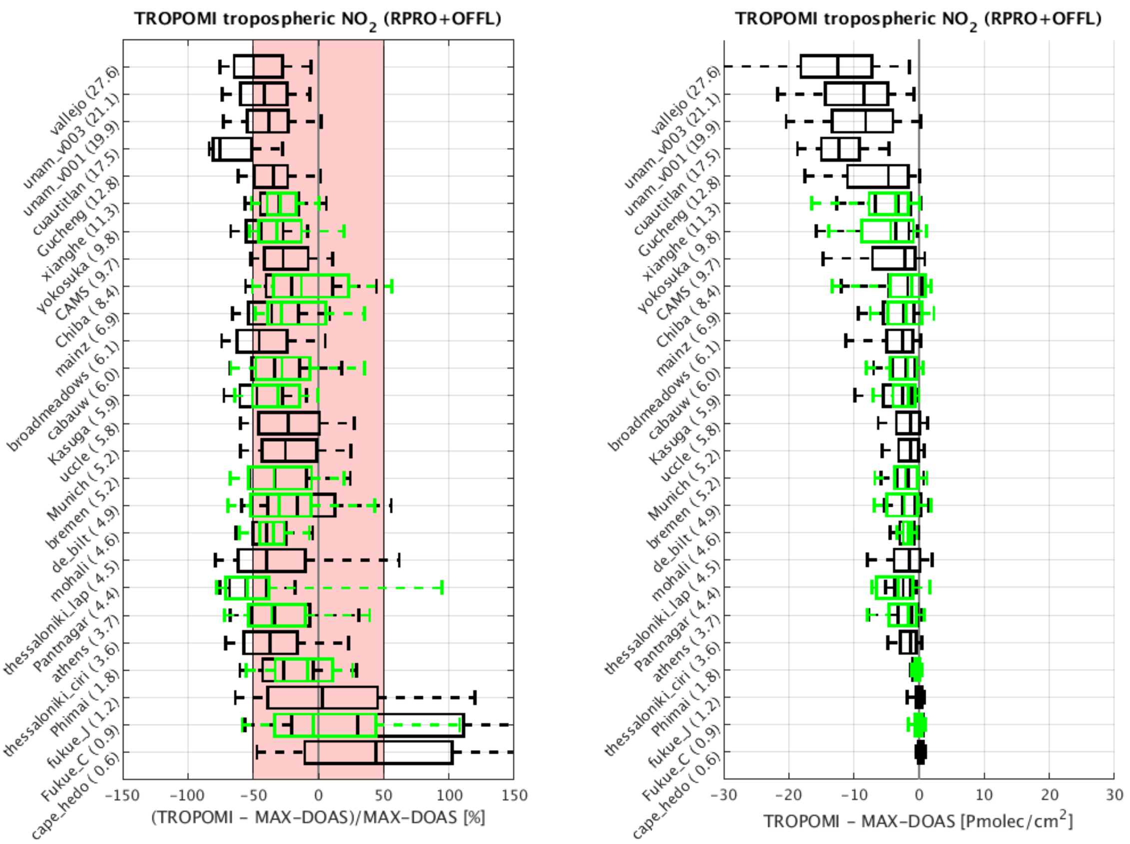


Figure 23: Box-and-whisker plots summarizing the bias and spread of the difference between S5P TROPOMI RPRO+OFFL and MAX-DOAS NO_2 tropospheric columns. Black corresponds to data up to and including processor version 1.3. Green corresponds to processor version 1.4, in operation since December 2020. Values between brackets in the labels denote the median tropospheric column at the station. The time frame is from May 2018 until May 2021. 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.

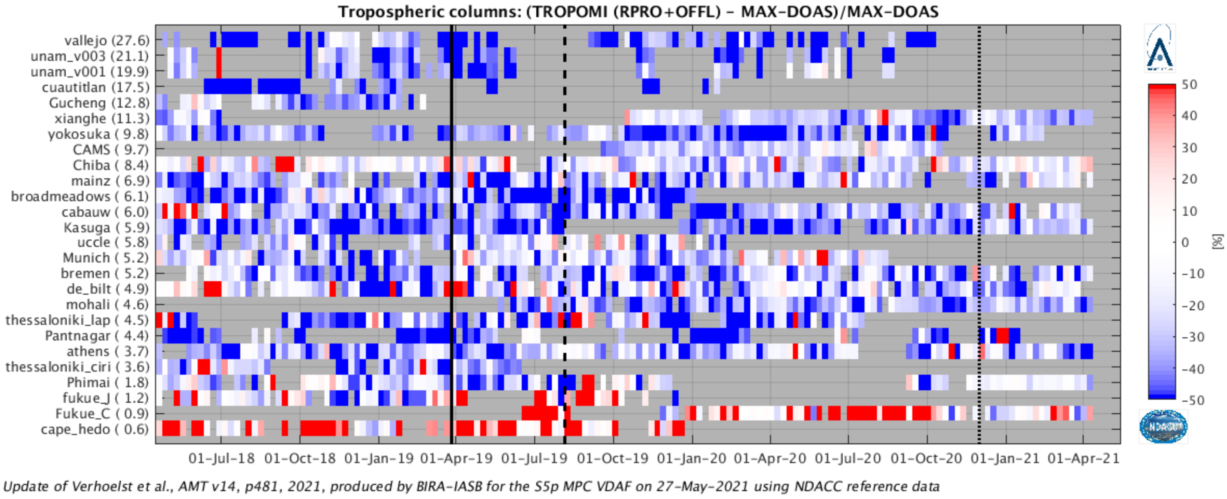


Figure 24: Time series – from May 2018 until mid of May 2021 – of the weekly averaged relative difference [%] between S5P RPRO+OFFL and MAX-DOAS NO_2 tropospheric column data. The black solid and dotted lines indicate, respectively, the processor switch from 1.2.2 to 1.3.0 and to 1.4.0, and dashed line the activation of the finer horizontal resolution. Stations are ordered by median tropospheric column.

A test case study has been performed for ground-based MAX-DOAS stations providing low tropospheric profiles, which is the case for the BIRA-IASB, ChibaU, JAMSTEC, and UNAM sites. 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:

$$VCD_{smoothed} = AK_{sat} * x_{MAXDOAS}$$

This analysis is illustrated on the 12 sites shown in **Figure 25** 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 (for 8 cases over 12 sites) by about 10 to 20% (see bottom of **Figure 25**). 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 and Cape Hedo can be attributed to the relative bias calculation from dividing with small tropospheric VCDs at very clean ground-based stations.

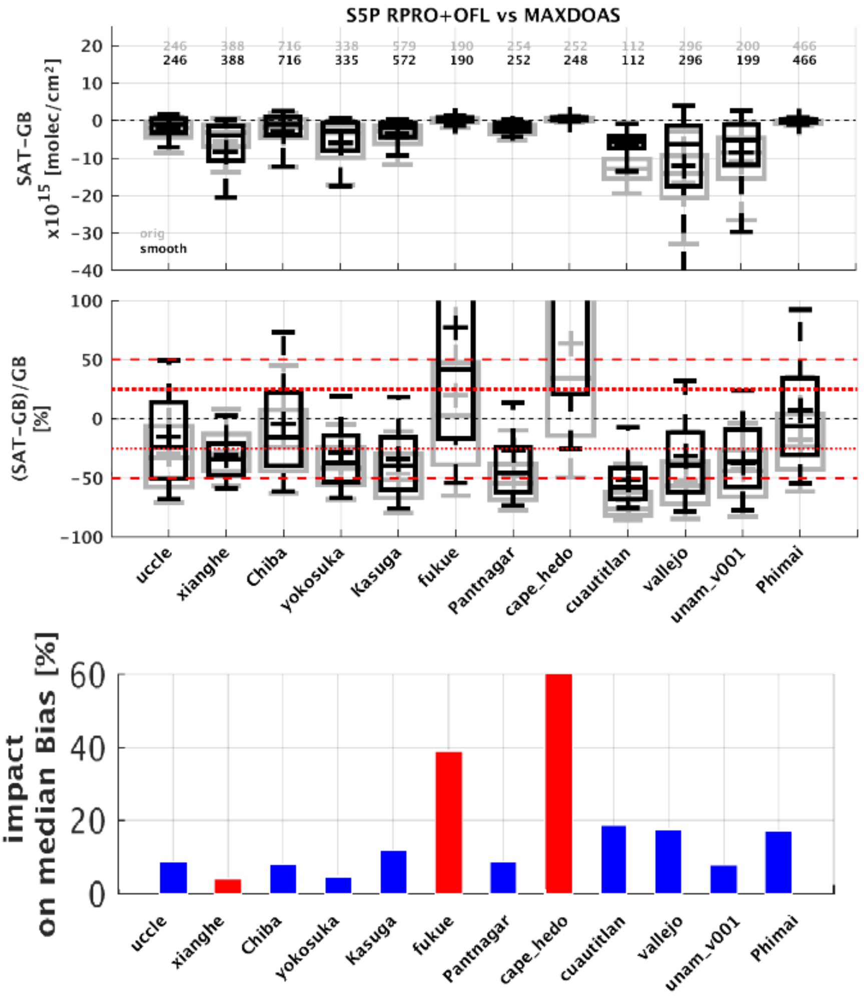


Figure 25: Box-and-whisker plots summarizing the bias and spread of the difference between S5P TROPOMI RPRO+OFFL and MAX-DOAS NO_2 tropospheric columns (grey) and the smoothed MAX-DOAS NO_2 tropospheric columns (black). The third panel presents bar plot summarizing the change on the median bias between the original comparisons and the smoothed ones. Absolute values of the relative median biases are shown, where reductions are marked in blue bars and increases in red bars. The time frame is from May 2018 until mid-February 2021. 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 and the crosses denotes the mean differences.

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%.

6.3.3.2 Dispersion

The median dispersion with respect to NIDFORVAL MAX-DOAS sites is 3 Pmolec/cm^2. Using again the three site types identified in the previous section (low polluted < 2 Pmolec/cm^2, medium polluted 3-14 Pmolec/cm^2, and extremely polluted > 15 Pmolec/cm^2), the median IP68/2 dispersion of the differences is about 0.7 Pmolec/cm^2, 3 Pmolec/cm^2, and 7.3 Pmolec/cm^2, respectively. The uncertainty precision requirement of maximum 0.7 Pmolec/cm^2 is only satisfied for the clean-sites ensemble.

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 spread (IP68/2) around the OLS regression line instead of the spread between S5P and MAX-DOAS data (adapted from Schneider et al., 2006). The residual spread (IP68/2) at the different European AVS sites is approximately 1.5 Pmolec/cm^2 and 2.5 Pmolec/cm^2 at Xianghe. There is a reasonably good correlation between TROPOMI and MAX-DOAS tropospheric columns with the Pearson R varying between 0.59 (Bremen) and 0.82 (Xianghe) with a mean of 0.74.

6.3.3.3 *Dependence on influence quantities*

Two key influence quantities for observations of tropospheric NO_2 are aerosol optical depth (AOD) and satellite cloud (radiance) fraction (CRF). The first one is retrieved within the MAX-DOAS NO_2 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 26**).

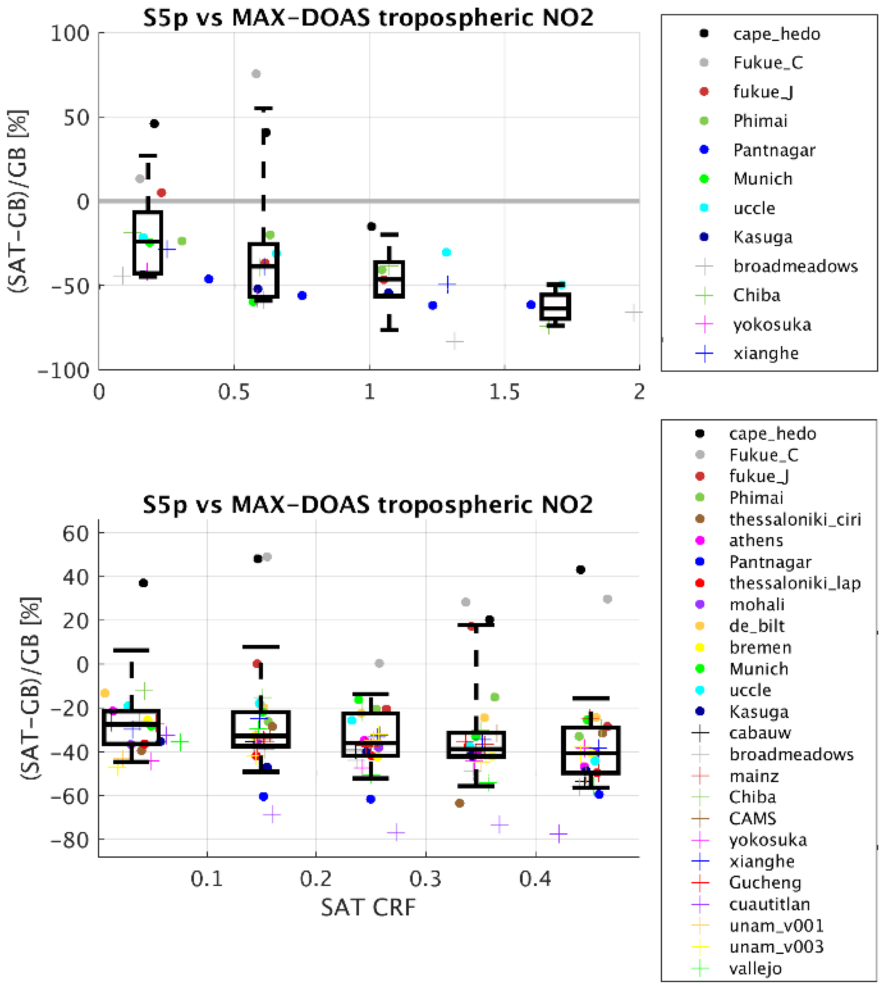


Figure 26: Relative differences between S5p TROPOMI RPRO+OFFL and MAX-DOAS NO_2 tropospheric columns as a function (upper panel) of MAX-DOAS AOD and (lower panel) S5p cloud radiance fraction (CRF). The time frame is from May 2018 until May 2021.

6.3.3.4 Seasonal and shorter term variability

The zonal daily mean NO_2 tropospheric columns for full mission lifetime are shown in **Figure 27**7. After the V01.04.00 change, an small increase is detectable in the Northern hemisphere between 30° and 60°N. Overall, no trends are visible.

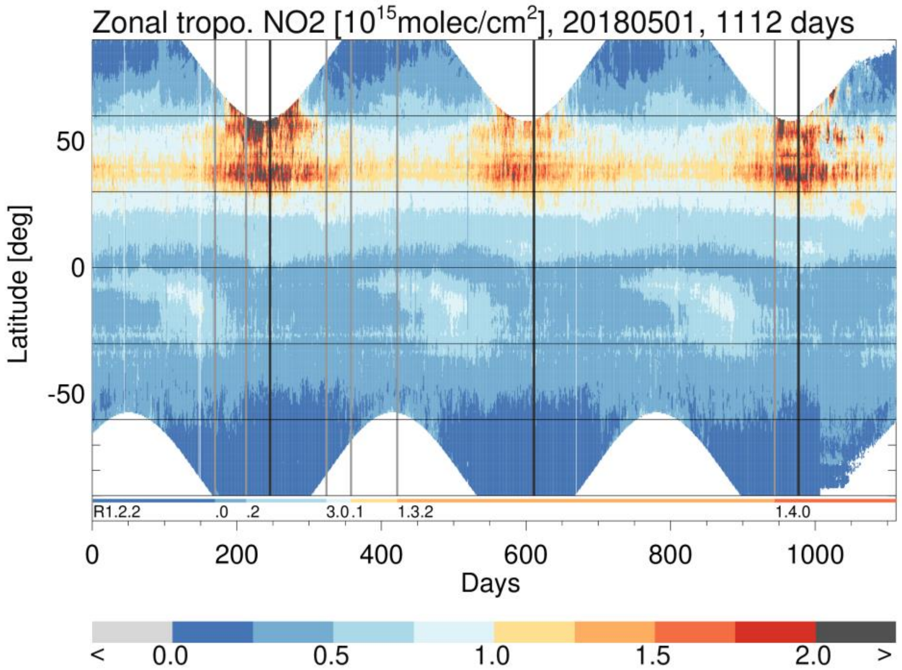


Figure 27: S5P TROPOMI NO_2 tropospheric columns [10^{12} molec/cm^2] as a function of day and latitude. The period is from May 2018 until mid-May 2021. Grid box size in latitude direction is 0.5°. The gray vertical lines mark the processor version changes, the black lines the beginning of each year.

Figure 28 presents the seasonal cycle of differences between S5P RPRO+OFFL and NDACC MAX-DOAS tropospheric NO_2. 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. TROPMI measures lower values than MAX-DOAS in late fall and in wintertime, 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.

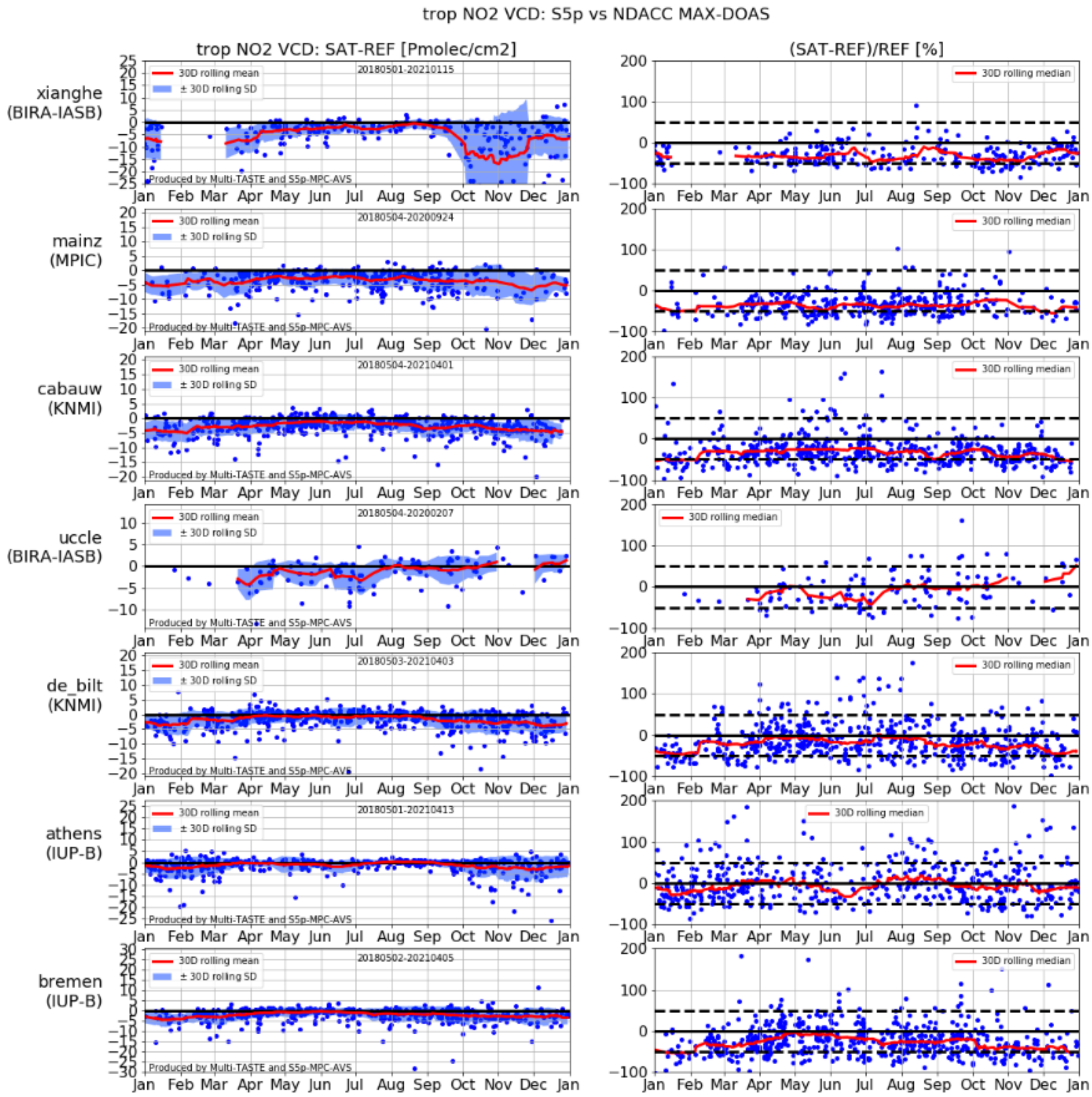


Figure 28: Seasonal cycle (with data mapped to one generic year) of the difference between S5P RPRO+OFFL and MAX-DOAS NO_2 tropospheric column data at six European stations. Difference (left) and relative difference (right). Data was obtained from the VDAF Automated Validation Server on 2021/05/27.

6.3.3.5 Geographical patterns

In general, no geographical patterns or artefacts can be detected in the latest L2_NO2 OFFL versions, as shown in the monthly mean over central Europe (**Figure**). A column increase is found in some areas (e.g. Warsaw) in V1.4.0 due to the retrieval changes of cloud pressure but more measurements are excluded due to quality filtering. The COVID related reductions came after January 2020 and do not play a role in these comparisons.

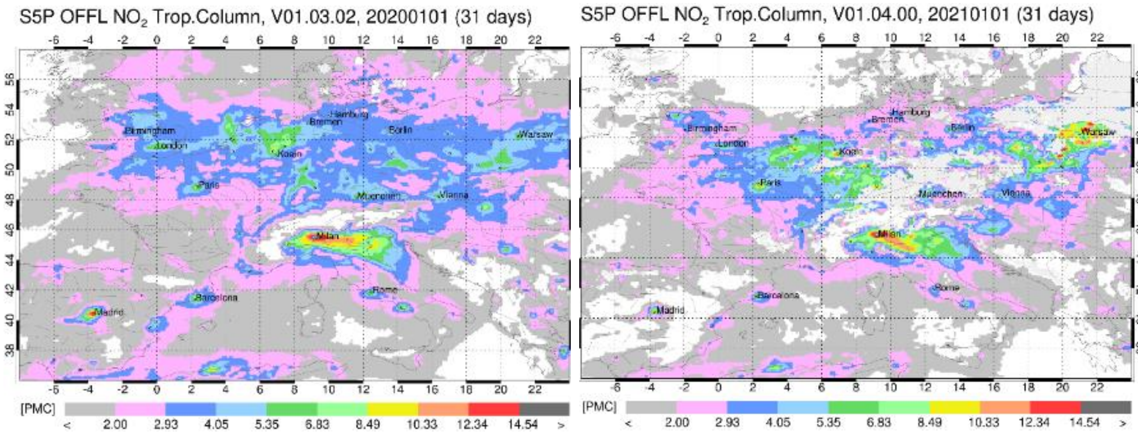


Figure 29: S5P OFFL tropospheric NO_2 over central Europe. The data is binned on a 0.1°x0.1° grid for January 2020 (V1.3.2) and 2021 (1.4.0). A `qa_value` > 0.75 and a cloud fraction `CF`<0.6 are used to reduce the amount of data and exclude cloudy scenes. 1 PMC means 1 Pmolec/cm^2.

6.3.3.6 Processor change from version 1.3 to version 1.4.0

Changes related to the version change v1.4 are clearly detectable for some sites (**Figure 23**). 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 2021 (v1.4), 2020 (v1.3), and 2019 (mix of v1.2 and v1.3). The results are given in the **Table 5**. The v1.4 change leads to a reduction of about 10% that started in December 2020.

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

	JFM 2019	JFM 2020	JFM 2021
Collocation count	784	714	623
Median bias [%]	-38.6	-38.1	-27.7
Dispersion (0.5 IQ68) [Pmolec/cm^2]	5.2	4.4	3.3

Comparisons of S5P NO_2 v1.2 and v1.3 with the QA4ECV OMI NO_2 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, where OMI uses the O2-O2 absorption feature and TROPOMI the O-2A absorption and the 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.

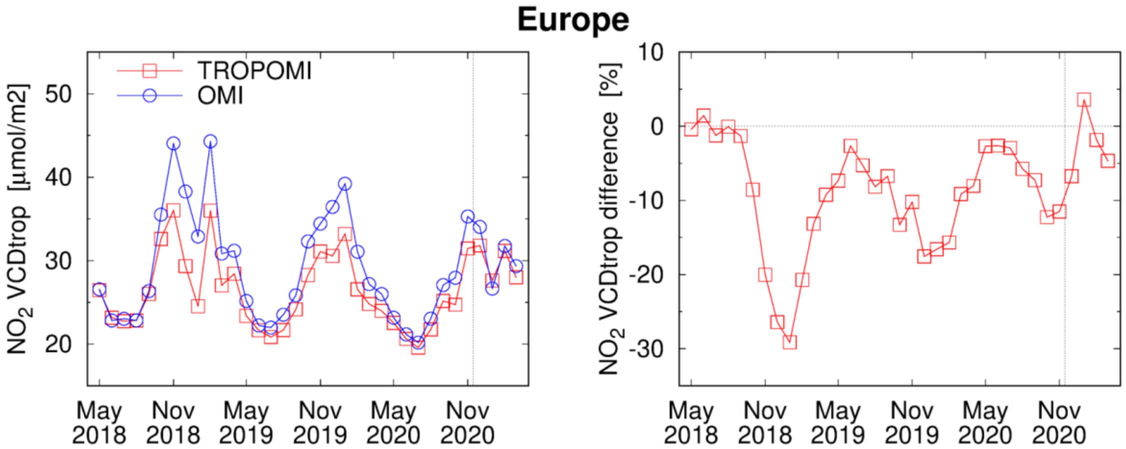


Figure 29: TROPOMI and OMI-QA4ECV NO_2 time evolution of the tropospheric NO_2 column (VCDtrop) for Europe [10W-25E; 35N-60N] in absolute numbers and as relative differences. TROPOMI data for pixels with `qa_value` > 0.75 and OMI/QA4ECV data with cloud radiance fraction < 0.5 (i.e. "clear-sky pixels") have been gridded to a common grid of 0.8°lon x 0.4°lat and are averaged over calendar months. The relative difference is defined with respect to the average of the two.

The tropospheric column shows a very good agreement between TROPOMI and OMI. This is demonstrated in **Figure 29** for Europe as an example. In general, TROPOMI columns are about 3% higher than OMI-QA4ECV (for Europe) and show a very similar seasonal cycle. The small systematic difference can be related to details of the settings of the DOAS retrieval. The tropospheric column comparisons for TROPOMI v1.3 and v1.4 retrievals are shown for three regions in **Figure 30** (Eastern China, Eastern USA, and Middle East). For polluted regions and hotspots, the tropospheric column of TROPOMI is lower than OMI, with negative differences varying from a few % up to -40%. The largest differences occur in winter.
Despite this systematic low bias compared to OMI, the correlation between the two datasets, both re-gridded to a 0.8° by 0.4° grid, is high, showing that OMI and TROPOMI are observing the same spatial variability of the concentrations (see also **Figure 31**). Note that representativeness differences will remain. Since the OMI ground pixels are much larger than those of TROPOMI, in particular near the edges of the swath, the re-gridded data will show different amounts of NO_2 and cloud cover. In addition, OMI data suffers from the row anomaly which affects the spatial sampling.

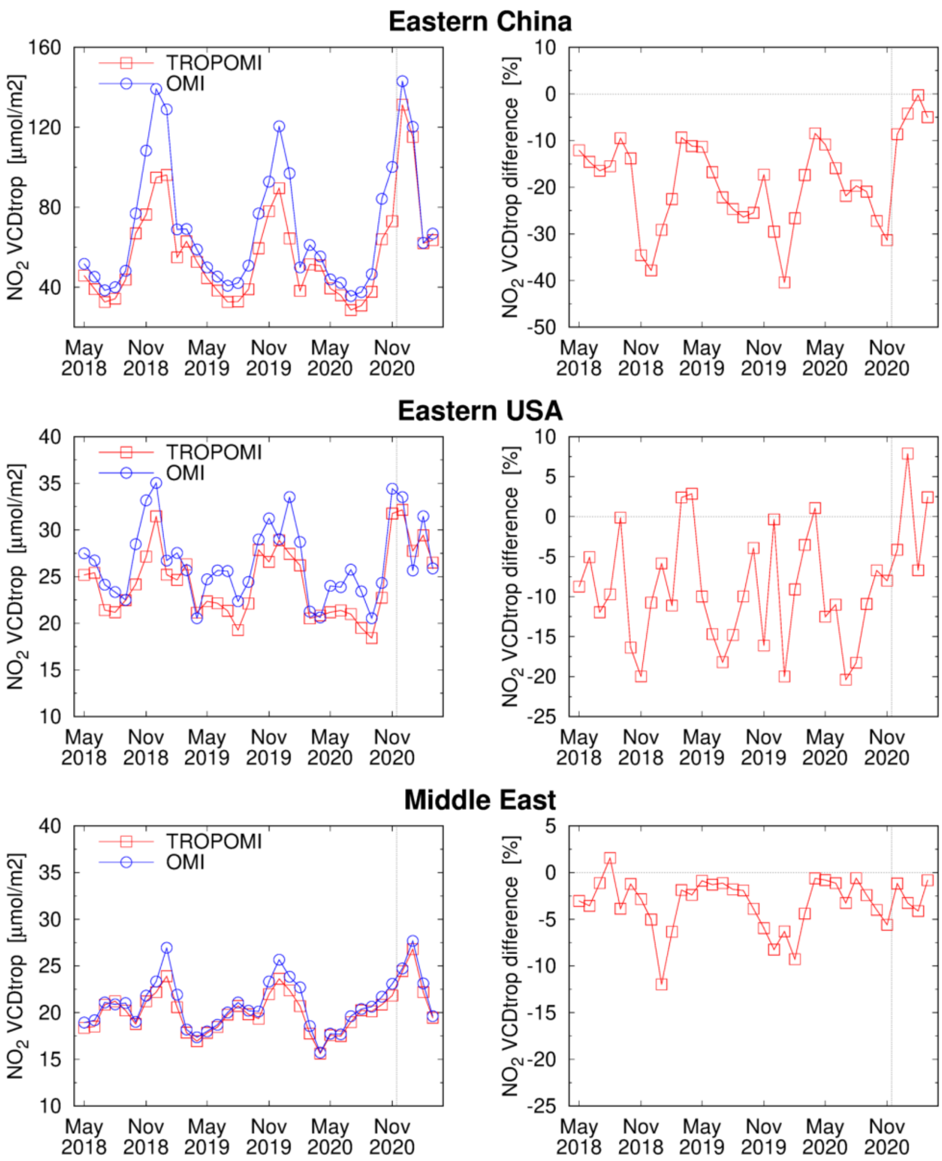


Figure 30: TROPOMI versions 1.2/1.3/1.4 and OMI-QA4ECV NO_2 time evolution of the tropospheric NO_2 column (VCDtrop) in absolute numbers (left) and as relative differences (right) for three regions: Eastern China [110E-124E; 21N-43N], Eastern USA [89W-69W; 32N-48N], and Middle East [30E-60E; 15N-40N]. 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 averaged over calendar months. The grey vertical line marks the beginning of V1.4.

On 2 December (NRTI) and 29 November (OFFL) 2020, the S5P NO_2 product was upgraded from v1.3.2 to v1.4.0. Comparisons of S5P NO_2 v1.4 with the QA4ECV OMI NO_2 shows a much-improved consistency between the two retrievals. In version 1.4, the FRESCO-S cloud retrieval was upgraded and now includes 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.

The impact of the upgrade is clearly demonstrated in **Table 6**. Substantial negative offsets compared to OMI are found for versions 1.2 and 1.3, with the largest relative differences for Eastern China in winter. For v1.4 the OMI and TROPOMI products are much more consistent, generally the regional monthly-mean observations agree to within 10%. The large difference over China in v1.3 (January 2020), and the improvement of v1.4 (January 2021) is also demonstrated in **Figure 31** and **Figure 32**. For v1.4, the comparison with OMI shows very similar values of the average tropospheric NO_2 column, as well as the spatial distribution over China.

Table 6 - Monthly and regional mean relative differences between TROPOMI and OMI-QA4ECV tropospheric NO_2 for the winter period November to March. The green cells refer to TROPOMI v1.4.0, the others to TROPOMI versions 1.3.x and 1.2.x

Eastern China	Nov.	Dec.	Jan.	Feb.	Mar.
2018	–35	–38	–29	–23	–9
2019	–17	–30	–40	–27	–17
2020	–31	–9	–4	–0	–5

Europe	Nov.	Dec.	Jan.	Feb.	Mar.
2018	–20	–26	–29	–21	–13
2019	–10	–18	–17	–16	–9
2020	–11	–7	+4	–2	–5

Eastern USA	Nov.	Dec.	Jan.	Feb.	Mar.
2018	–20	–11	–6	–11	+2
2019	–16	–0	–20	–9	–4
2020	–8	–4	+8	–7	+2

Middle East	Nov.	Dec.	Jan.	Feb.	Mar.
2018	–3	–5	–12	–6	–2
2019	–6	–8	–6	–9	–4
2020	–6	–1	–3	–4	–1

India	Nov.	Dec.	Jan.	Feb.	Mar.
2018	–12	–11	–13	–11	–7
2019	–17	–14	–14	–12	–9
2020	–13	–9	–11	–6	–2

Central Africa	Nov.	Dec.	Jan.	Feb.	Mar.
2018	+0	–4	–7	–6	–8
2019	–7	–6	–7	–4	–6
2020	–7	–7	–6	–5	–6

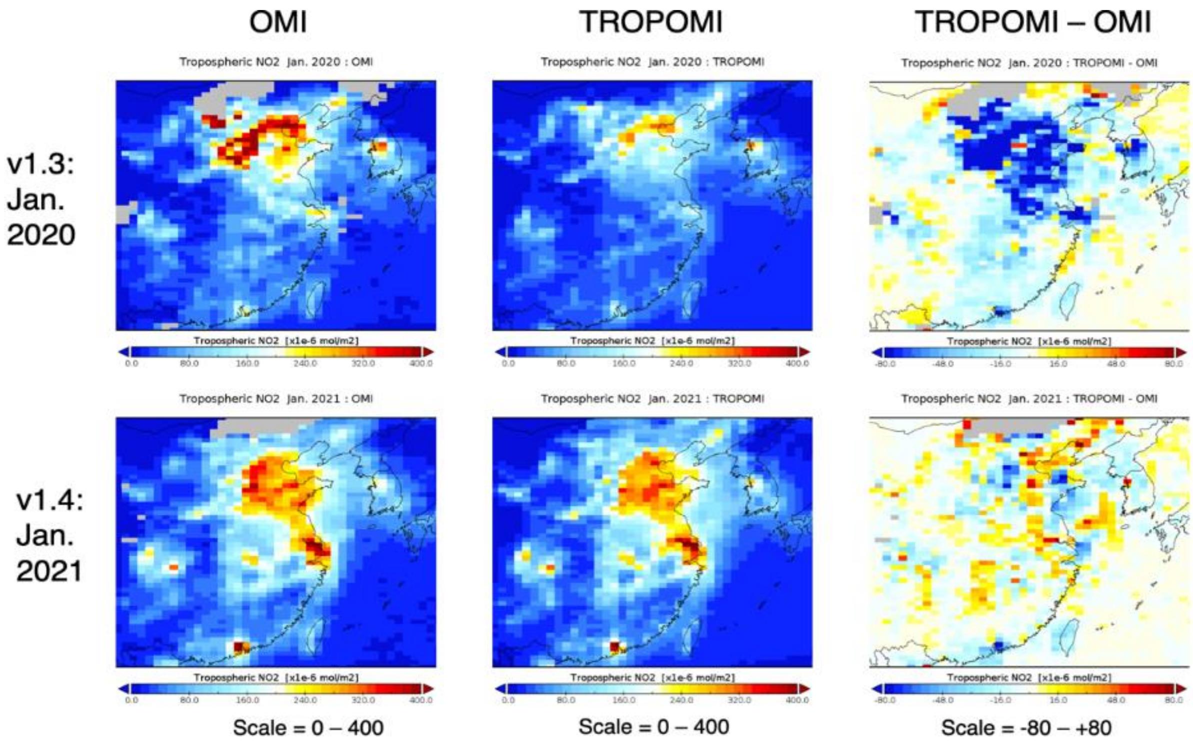


Figure 31: TROPOMI and OMI-QA4ECV NO_2 monthly-mean maps for Eastern China [110E-124E;21N-43N]. Shown are the tropospheric NO_2 column (VCDtrop) of OMI (left), TROPOMI (middle) and the difference (right) for January 2020 (top; v1.3) and January 2021 (bottom; v1.4). TROPOMI data for pixels with `qa_value` > 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 averaged over calendar months.

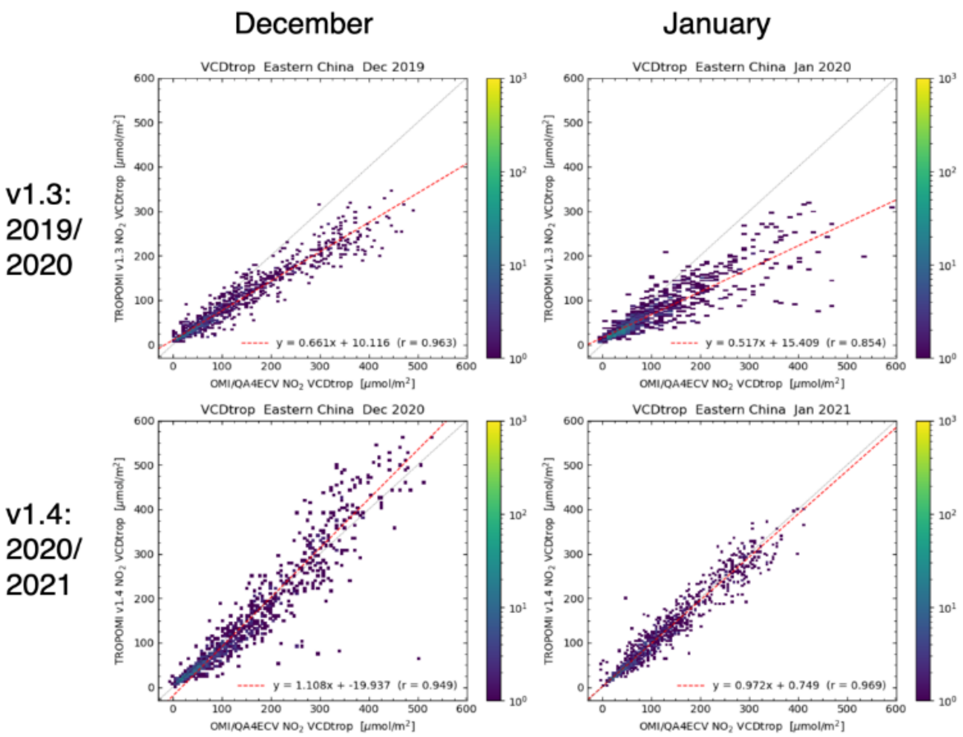


Figure 32: TROPOMI-OMI tropospheric NO_2 monthly scatter plots for Eastern China [110E-124E;21N-43N]. Shown are the tropospheric NO_2 column (VCDtrop) of TROPOMI (y-axis) versus OMI (x-axis) for December and January 2019/2020 (top; v1.3) and 2020/2021 (bottom; v1.4). TROPOMI data for pixels with `qa_value` > 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.

6.3.3.7 Other features

Ordinary linear regression (OLS) of S5P vs MAX-DOAS yields fairly good correlation coefficients (0.59-0.84, see before), but low slopes. When $S5P=a*MXD+b$, a varies between 0.3 (Bremen, Uccle) to 0.6 (Athens). It is known, however, that this approach is only correct in the limit that all random errors are in S5P data. For the MAX-DOAS vs S5P OLS (i.e., assuming the opposite limit that all random errors are in MAX-DOAS), one obtains slopes closer to unity: a varies then between 0.71 (Mainz) and 0.99 (Uccle).

6.3.4 Stratospheric NO_2 column

6.3.4.1 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 33** and **Figure 33**). 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 VDAF-AVS ZSL-DOAS stations (8836 co-locations), the OFFL stratospheric NO_2 column mean bias is -5.4% (-0.15 Pmolec/cm^2) with a station-to-station 1σ scatter of 9%. 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).

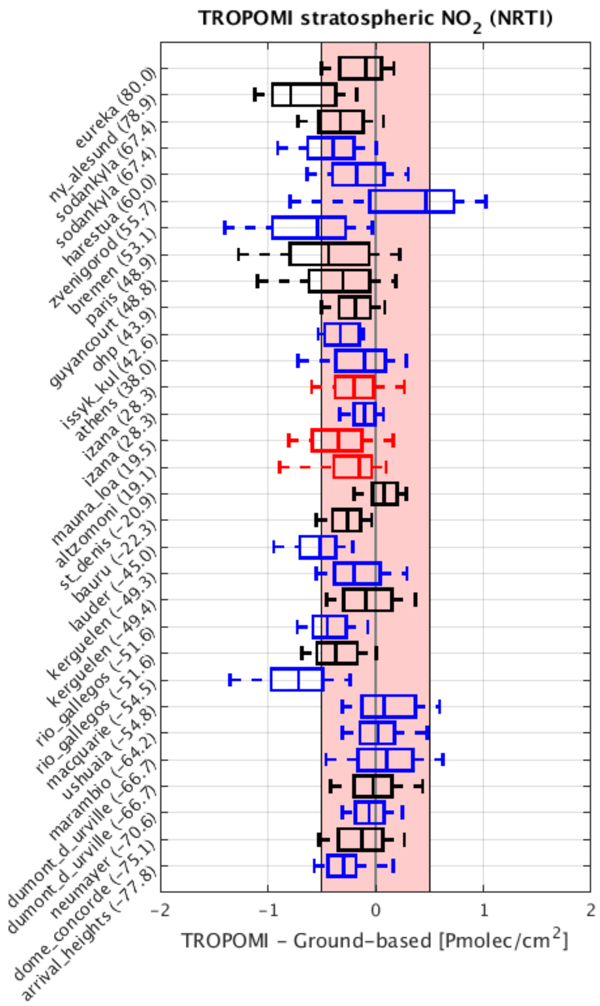


Figure 33: Pole to pole bias and spread [Pmolec/cm^2] of S5P TROPOMI NRTI and NDACC ZSL-DOAS NO2 stratospheric columns (SAOZ data in black, other ZSL-DOAS in blue, and mountain-top PGN in red). The network-wide mean bias and its formal uncertainty are -0.17 and 0.06 Pmolec/cm^2, respectively. Stations 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 time frame is May 2018 to May 2021. 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.

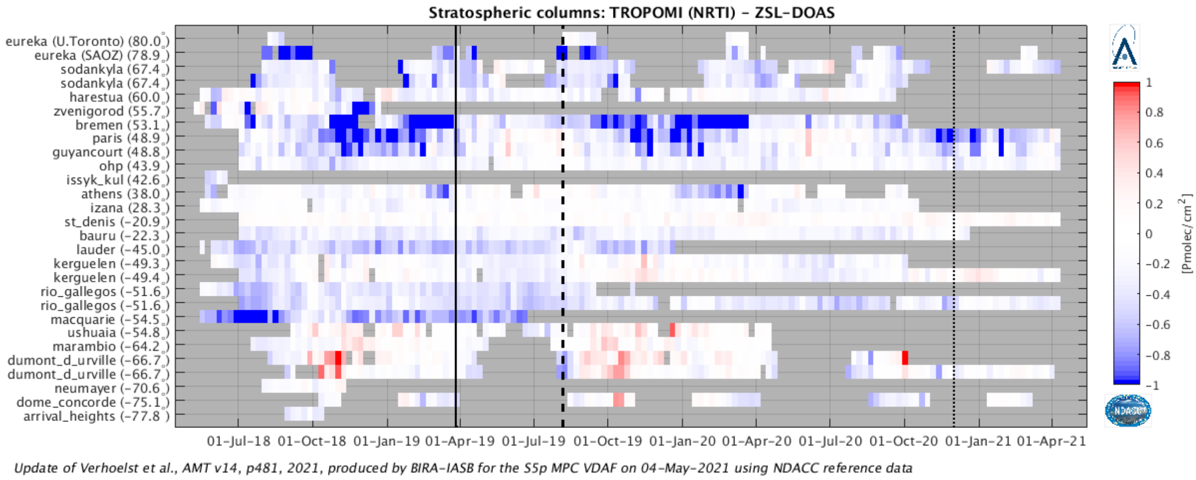


Figure 34: Time series – from August 2018 until May 2021 – of S5P NRTI L2_NO2 and NDACC ZSL-DOAS NO_2 stratospheric column differences, weekly averaged [Pmolec/cm^2]. The black solid, dotted, and dashed lines indicate, respectively, the processor switches (from 1.2.2 to 1.3.0 and from 1.3.2 to 1.4.0), and 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.

6.3.4.2 Dispersion

The ±1σ dispersion of the difference between stratospheric column and reference data around their median value rarely exceeds 0.3 Pmolec/cm² at sites without tropospheric pollution (cf. the error bars in **Figure 33**). When combining random errors in the satellite and reference measurements with irreducible co-location 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.

6.3.4.3 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 S5P measurement does not reveal bias variations much larger than 0.4 Pmolec/cm^2 over the range of the influence quantities (**Figure 35**).

6.3.4.4 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 36. 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.

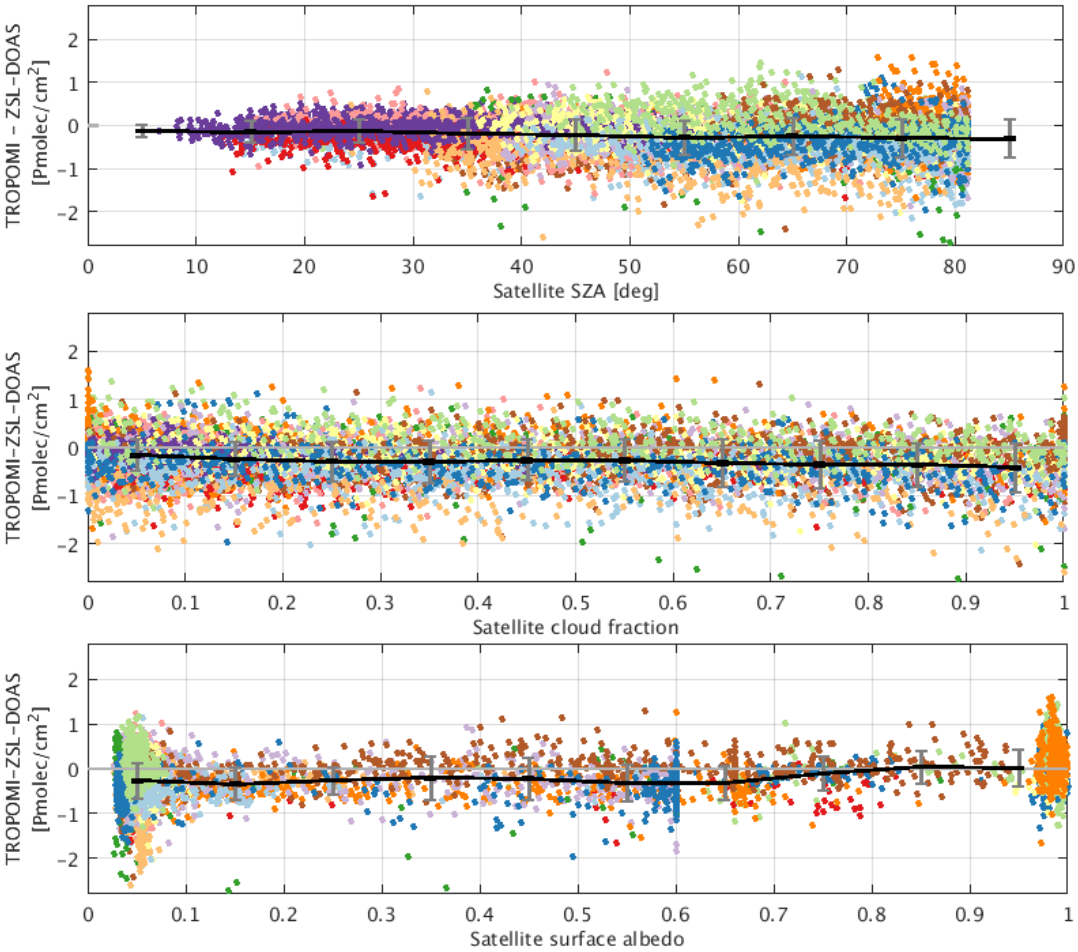


Figure 35: Difference between S5P L2_NO2 NRTI and ground-based SAOZ stratospheric NO_2 columns as a function of the satellite solar zenith angle (SZA), satellite cloud fraction, and satellite surface albedo. Mean and standard deviation calculated over bin widths of 10° degrees in SZA, 0.1 in CF, and 0.1 in surface albedo (solid black line and grey bars). Co-locations cover the period May 2018 – February 2021.

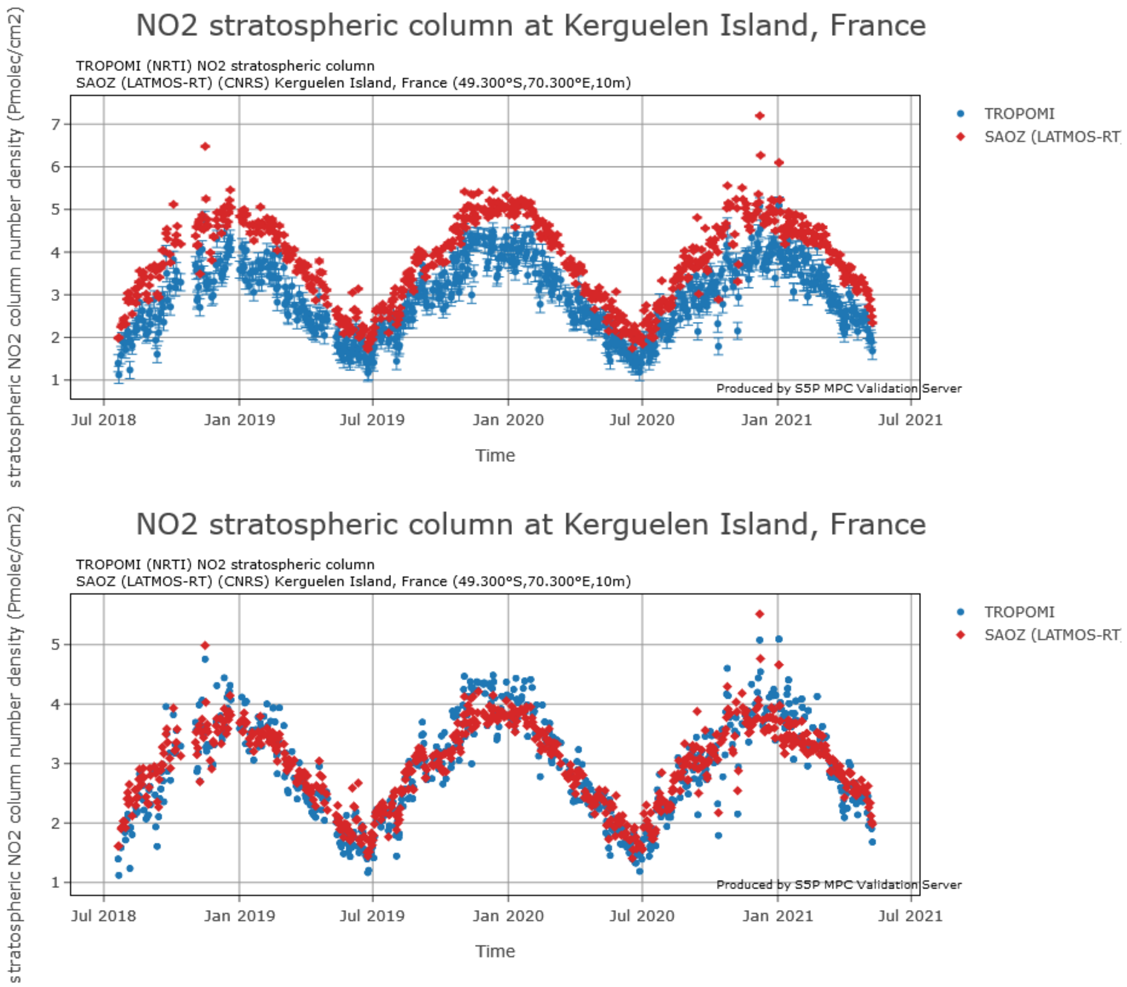


Figure 36: Time series of S5P NRTI L2_NO2 stratospheric NO2 column data (blue dots) co-located with ground-based SAOZ twilight measurements (red dots) at sunset performed by LATMOS at the NDACC southern mid-latitude station of Kerguelen Island. In the upper plot, the photochemical correction is deactivated to offset the two time-series and to better see the day-to-day variability. The time frame is July 2018 to May 2021.

TROPOMI stratospheric NO_2 and Pandora total NO_2 at three mountain sites plotted follow the same seasonal cycle (**Figure 37**). 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.

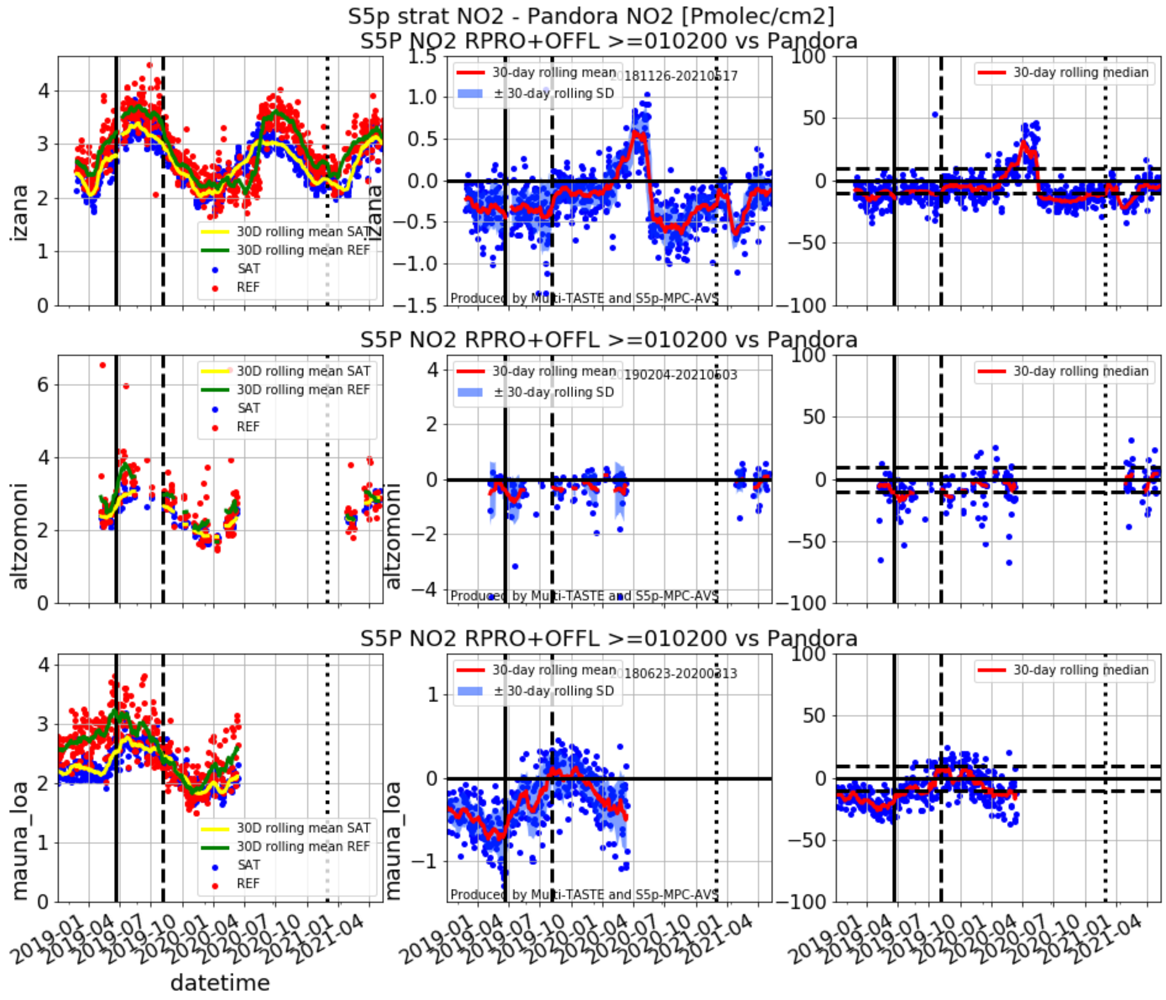


Figure 37: Time series of the S5P RPRO+OFFL L2_NO2 stratospheric NO_2 column and Pandora total column at 3 mountain sites. Left row: S5P and PGN Pandora together; centre: their difference; right: their percent relative difference. Plain lines represent the 30-day rolling mean or (for relative difference) the 30-day rolling median. Full vertical line: OFFL processor version change 1.2 to 1.3 on 2019/03/20. Dashed line: pixel size switch at 2019/08/06. Dotted line: OFFL processor version change 1.3 to 1.4 on 2020/11/29. Data covers the period June 2018 until May 2021. It is obtained from the validation server on 2021/05/27.

6.3.4.5 Geographical patterns

None to report.

6.3.4.6 Processor version change 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), 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 34) and 3 PGN Pandora mountain sites (Figure 37). 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 is expected to affect mainly the tropospheric part of the column.

6.3.4.7 Other features

None to report.

6.3.5 Total NO_2 column

6.3.5.1 Bias

Based on measurements from 35 Pandora stations between 78.9°N and -45.8°S available at VDAF-AVS (**Figure 38**), the median bias (15176 co-locations, checked 2021/05/24) is -6.7% (-0.5 Pmolec/cm^2) with a high station-to-station scatter of 18%. 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/cm^2), the bias is +4.1% for the 21 less polluted stations and -22.6% for 14 high polluted stations.

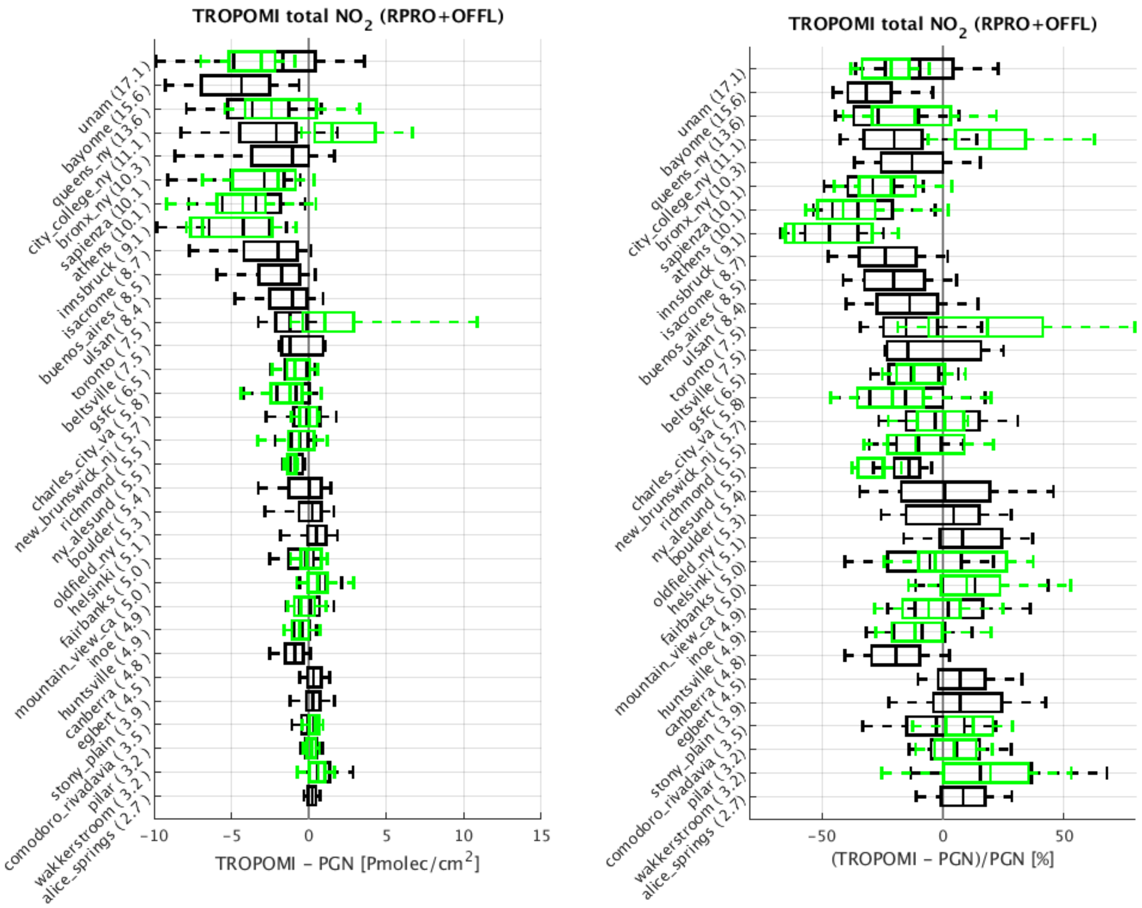


Figure 38: Box-and-whisker plots summarizing the relative bias and spread [%] between S5P TROPOMI RPRO+OFFL (processor version 1.2.2 up to 1.3.2 in black, version 1.4.0 in green) and PGN Pandora NO_2 total column data. Conventions of the boxplots are identical to Fig. 21. Stations are ordered by median total column. The time frame is May 2018 to May 2021. Note that the number of co-locations for v1.4 is still limited, hence the somewhat unstable determination of the whiskers at some sites.

We highlight three different comparison cases here. At Alice Springs (Australia), where the total NO_2 column values are mostly between 2-4 Pmolec/cm^2, a small positive mean bias of 0.2 Pmolec/cm^2 is seen (+8% median bias). A wider distribution of NO_2 values (2-30 Pmolec/cm^2) is found at Bronx (New York, United States) where the bias is -3 Pmolec/cm^2 (-13%). Finally, at Sapienza (Rome, Italy) where column values can reach up to 40 Pmolec/cm^2, the bias is -4 Pmolec/cm^2 (-29%), probably due to locally enhanced NO_2. The bias has been reduced in the months after the V1.4.0 change.

6.3.5.2 Dispersion

The dispersion of the S5P and PGN Pandora differences depends strongly on the station. Small uncertainties (IP68/2) are observed at Ny-Alesund, Alice Springs, Mauna Loa, Pilar, Izana, Comodoro Rivadavia (0.4-0.6 Pmolec/cm^2) that are within the mission precision requirement, and higher values elsewhere (e.g., 3-5 Pmolec/cm^2 at New York Bronx, Sapienza Rome, New York City College, Unam). The per-site median of the IP68/2 is 1.5 Pmolec/cm^2. The Pearson-R is 0.65 and varies from relatively low (e.g., 0.33 at Charles City) to high (0.88 at Wakkerstrom).

6.3.5.3 Dependence on influence quantities

None to report.

6.3.5.4 Short term variability

None to report.

6.3.5.5 Geographical patterns

None to report.

6.3.5.6 Processor version change and switch to smaller ground pixel resolution

Two major processor version changes (1.2 to 1.3 to 1.4) occurred for OFFL L2_NO2 on 20 March 2019 and 2 December 2020. 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 39** 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 sites), as expected from the improvements in the cloud parameters.

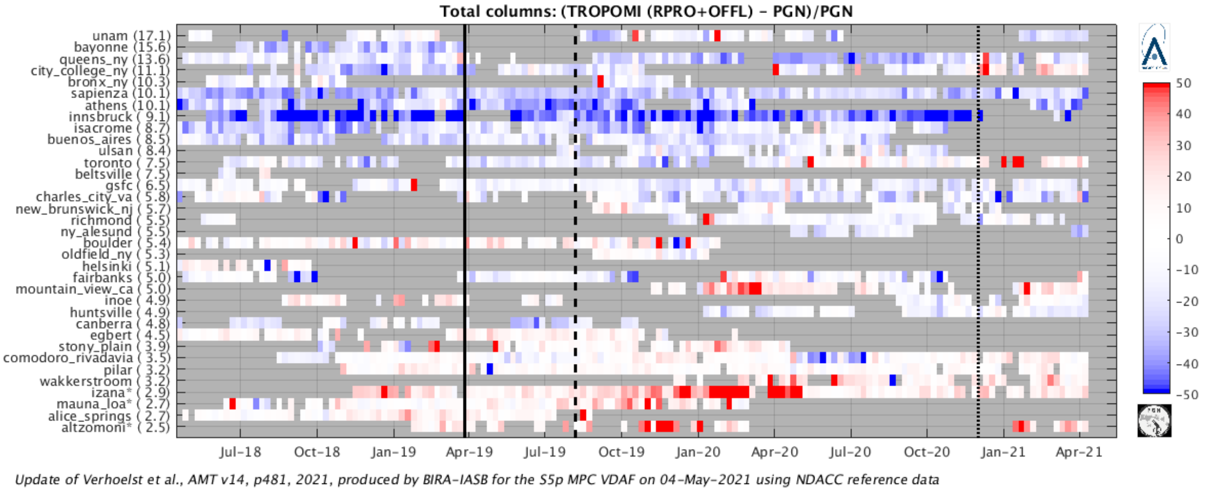


Figure 39: Time series – from May 2018 until May 2021 – of S5P RPRO+OFFL and PGN Pandora NO_2 total column differences, weekly averaged [%]. The black solid and dashed lines indicate, respectively, the processor switch from 1.2.2 to 1.3.0 and the activation of the finer horizontal resolution; the dotted line the switch from 1.3.2 to 1.4.0. Stations are ordered by median total column.

Changes related to versions v1.4 are visible at some sites shown in **Figure 39**. To quantify this effect, we calculate the median relative difference and the difference dispersion for all the measurement pairs for the January-February-March period of 2021 (v1.4), 2020 (v1.3) and 2019 (mix of v1.2 and v1.3). The results are given in the **Table 7**. Bias and dispersion appear lower in 2021, but note the lower number of comparison pairs here, which could be due to less sites contributing data in this period, or a higher `qa_value` filter impact.

Table 7 - 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.

	JFM 2019	JFM 2020	JFM 2021
Collocation count	859	978	591
Median [%]	-5.9	-6.5	-2.3
Dispersion (0.5 IQ68) [Pmolec/cm^2]	2.6	2.3	1.8

6.3.5.7 Other features

None to report.

6.4 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).

6.4.1 Tropospheric NO2 Column NRTI vs OFFL

To demonstrate the closeness of the NRTI and OFFL L2_NO2 products at MAX-DOAS sites, L2_NO2 NRTI (processor version 01.00.02 to 01.04.00) and L2_NO2 OFFL (RPRO processor version 01.02.02 + OFFL processor version 01.02.00 to 01.04.00), 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 Athens are summarized in **Table 8**. Similar conclusions on the closeness of NRTI and OFFL can also be drawn for the other sites.

Table 8 - Statistics on the comparison of the common subset of L2_NO2 NRTI, L2_NO2 RPRO+OFFL and co-located MAX-DOAS, for the sites Bremen and Athens (*: unit of Pmolec/cm2).

Bremen: 428 common co-locations Orbits range from 3364 (2018-06-07) to 18019 (2021-04-05).			
	NRTI-OFFL	NRTI-MXD	OFFL-MXD
Mean(diff) ±sem*	0.07	-2.16±0.17	-2.24±0.17
Median(diff)*	0.06	-1.31	-1.34
Std(diff)*	0.4	3.5	3.5
1/2 IP68(diff)*	0.2	2.4	2.5
Pearson R	0.98	0.56	0.55
Slope	1.00	0.30	0.29
Athens: 421 common co-locations Orbits range from 3392 (2018-06-09) to 18132 (2021-04-13).			
	NRTI-OFFL	NRTI-MXD	OFFL-MXD
Mean(diff) ±sem*	0.11	-1.29±0.18	-1.40±0.18
Median(diff)*	0.09	-0.17	-0.24
Std(diff)*	0.2	3.6	3.6
1/2 IP68(diff)*	0.1	2.2	2.2
Pearson R	1.00	0.81	0.82
Slope	1.02	0.57	0.56

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.

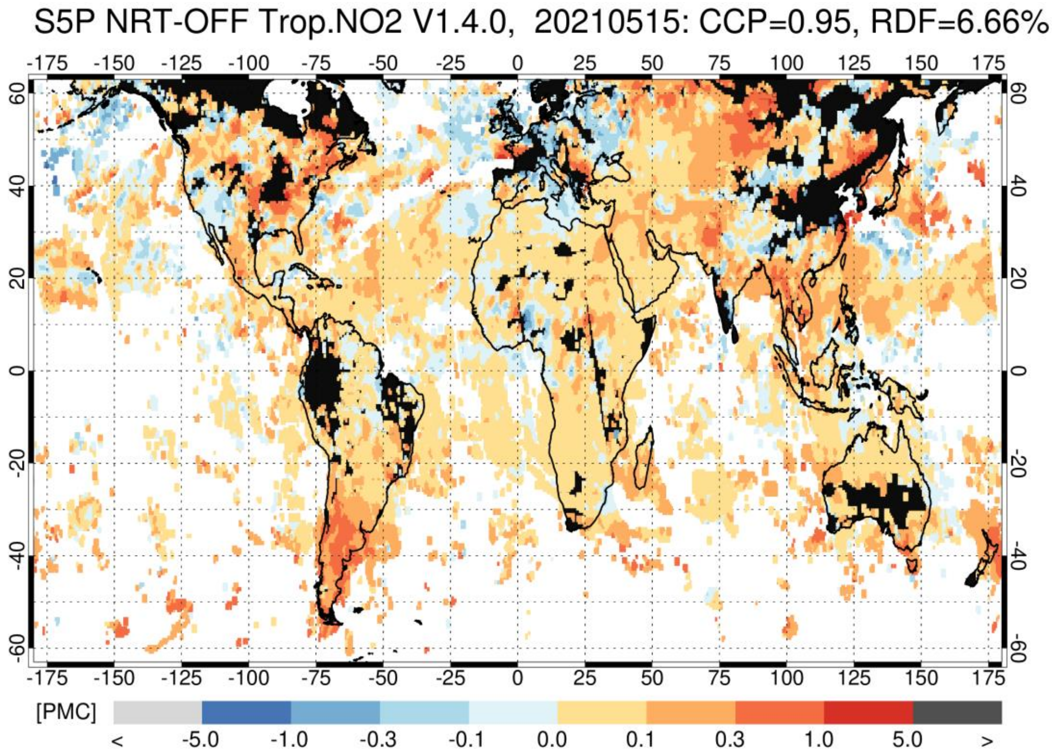


Figure 40: Difference between S5P NRTI and OFFL V1.4.0 daily tropospheric NO_2 column data (binned to 0.5°x0.5° resolution) for May 15, 2021. The difference is only calculated for columns above 0.5 Pmolec/cm^2.

Daily tropospheric columns from NRTI and OFFL V1.4.0 data streams are binned to 0.5° grid cells and differentiated (**Figure 40**) to elucidate the spatial distribution of differences. The Pearson correlation coefficient is high (0.95) and the global relative mean difference is below 7%. Differences above $\pm$1 Pmolec/cm^2 are rarely found.

6.4.2 Stratospheric NO_2 Column NRTI vs OFFL

The similarity of the two products can be investigated by comparing the processing of a randomly chosen orbit**. Figure 41** shows this approach for May 1st, 2021. It reveals differences mostly below the mission requirement on precision (0.5 Pmolec/cm^2). The RMSD is 0.14 Pmolec/cm^2, with values up to 0.5 Pmolec/cm^2. Some features are due to a different stratosphere/troposphere separation (e.g. over Greenland, positive difference in stratosphere, negative in troposphere, **Figure 41**). 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.

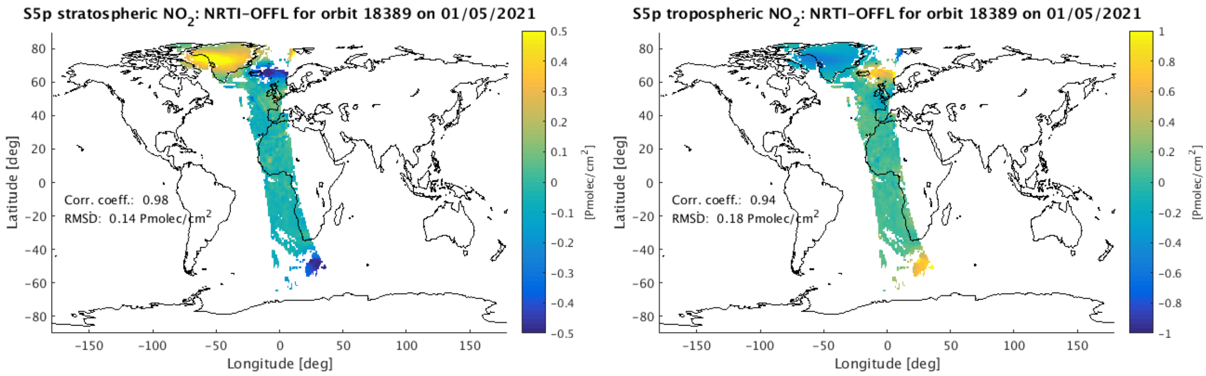


Figure 41: Difference between S5P NRTI and OFFL stratospheric (left panel) and tropospheric (right panel) NO_2 column data for a single orbit on May 1st, 2021. Data is gridded to 1°x1° resolution with a `qa_value` > 0.75.

A comparison of stratospheric NRTI and OFFL pole-to-pole validation graphs as shown in **Figure 42** illustrates that in direct validation studies the NRTI performs very similarly to OFFL with a median bias of less than -0.2 Pmolec/cm^2 at sunset.

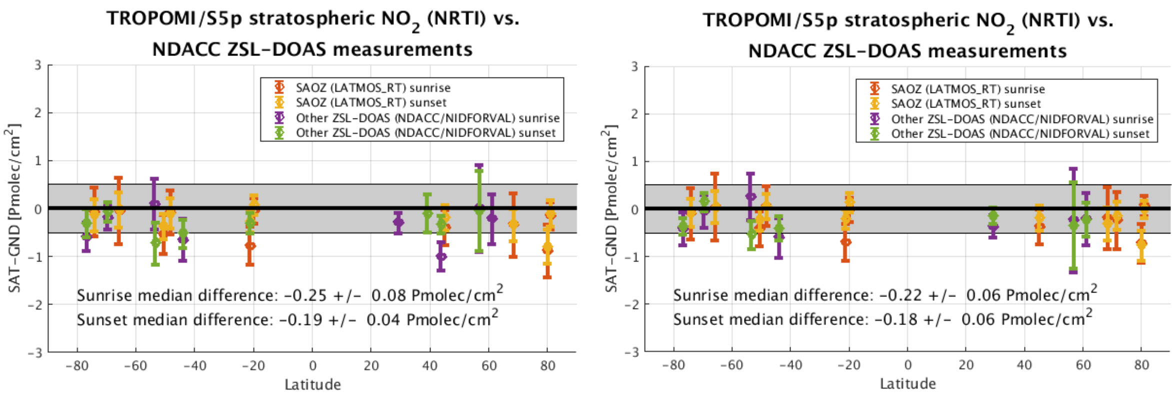


Figure 42: Meridian dependence of the mean (the circular markers) and dispersion (±1σ error bars) of the differences between S5P TROPOMI L2_NO2 (NRTI in the left panel, OFFL in the right) stratospheric column data and ZSL-DOAS reference data, represented at individual stations from the Antarctic to the 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 of phase E2 until May 2021.

6.4.3 Total NO_2 Column NRTI vs OFFL

The global relative total column difference is in the range of 0.8%, with NRTI being slightly higher. A single orbit 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 43**).

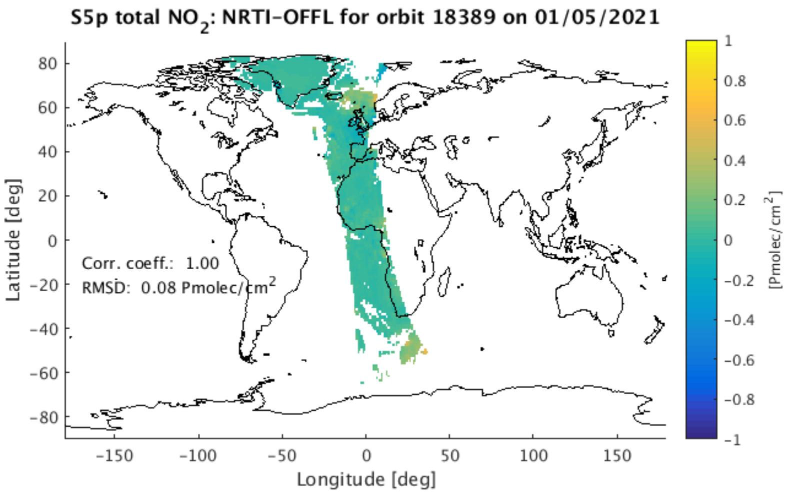


Figure 43: Difference between S5P NRTI and OFFL total NO_2 column data for a single orbit on May 1st, 2021. Data is gridded to 1°x1° resolution with a `qa_value` > 0.75.

6.5 Internal consistency of the validation results

This section focuses on the internal consistency checks of the NO_2 validation results. This rely on analysis of potential source of inconsistencies (such as impact of different NO_2 cross-sections) or synergistic analysis of several instruments types.

6.5.1 NO2 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 9**, 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 9 - NO2 cross section source and temperature for the different instruments (Verhoelst et al., 2021).

Instrument	reference	temperature	comments
S5p TROPOMI	Vandaele et al. (1998)	220K	With temperature correction in AMF (Zara et al., 2017)
ZSL-DOAS	Vandaele et al. (1998)	220K	
ZSL-DOAS	Harder et al. (1997)	227K	NIWA instruments
MAX-DOAS	Vandaele et al. (1996)	298K	tropospheric retrieval only
MAX-DOAS	Vandaele et al. (1998)	298K and 220K	Orthogonalized following Peters et al. (2017)
PGN	Vandaele et al. (1998)	254.4K	PGN processor v1.7

6.5.2 ZSL-DOAS and mountain-top PGN

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 sites, 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 sites 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/cm2 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 NO2. This needs to be confirmed with a future PGN release.

6.5.3 Sites with multiple instruments (different geometries)

There are a number of sites 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). Current sites with these two types of instruments are Athens, Thessaloniki, Xianghe, Uccle, 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 Validation Results: L2_HCHO

7.1 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 7.4 shows evidence that NRTI and OFFL data do not differ significantly and that their respective validations 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. The current L2_HCHO NRTI and OFFL product version 02.01.04 was released in November/December 2020.

7.2 Validation approach

7.2.1 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 S5PVT AO projects with faster data delivery (e.g., CESAR AO ID 28596 and NIDFORVAL AO ID 208607).

7.2.1.1 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/cm^2 (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:

- S5P 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 S5P 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.

7.2.1.2 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 pixel are kept for `qa_value` ≥ 0.5. It covers the MAX-DOAS measurement location, and is within ±0.5 h of the MAX-DOAS measurement. Only the closest MAX-DOAS measurement (in time) is used. No a priori substitution or averaging kernel is applied.

- NIDFORVAL AO project: S5P pixels with a `qa_value` ≥ 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.

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

The Xianghe site (China, 39.75° N, 116.96°E) is one of the few sites 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 site. Summary results (**Figure 45**) 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).

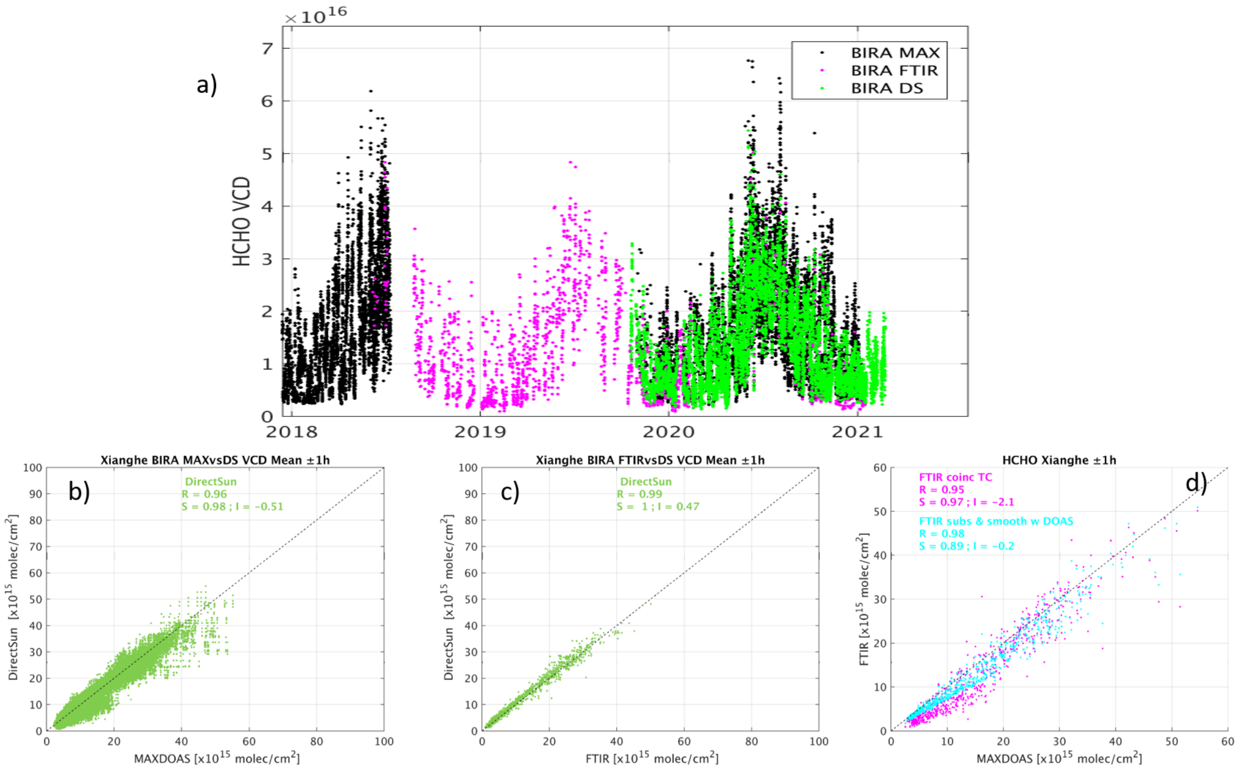


Figure 44: a)Time-series of the MAX-DOAS, FTIR and direct-sun HCHO dataset from BIRA in Xianghe; Scatter-plot between the b) direct-sun and the MAX-DOAS, c) direct-sun and FTIR and d) FTIR and MAX-DOAS for the raw comparisons and when taking into account a-priori profile substitution and smoothing.

7.2.2 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).

7.2.3 Field campaigns and modelling support

Nothing to report.

7.3 Validation of L2_ HCHO

7.3.1 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.

7.3.2 Status of validation

This section presents a summary of the key validation results obtained by the Validation Data Analysis Facility (VDAF) of the S5P Mission Performance Centre (MPC) and takes into account results obtained by S5P 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, five MAX-DOAS stations contribute data routinely to the VDAF Automated Validation Server with a temporal coverage of collocations from May 2018 to January 2021. Data from nine stations are used in the NIDFORVAL AO project. The period covers May 2018 to January 2021, depending 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 covering the period May 2018-January 2021.

7.3.3 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.

7.3.3.1 Fourier Transform Infrared Spectrometers

TROPOMI shows a positive bias of +27±4% for clean sites (HCHO <2.5 Pmolec/cm^2) and a negative bias of -29±1% for high emission sites (>8 Pmolec/cm^2) in comparison to correlative data from 28 NDACC FTIR stations, covering the period from May 2018 to December 2020, as illustrated in **Figure 45**. 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 1.04±0.04 Pmolec/cm^2 and a proportional bias of 0.65±0.032 is obtained (updated from Vigouroux et al. (2020), Fig. 4).

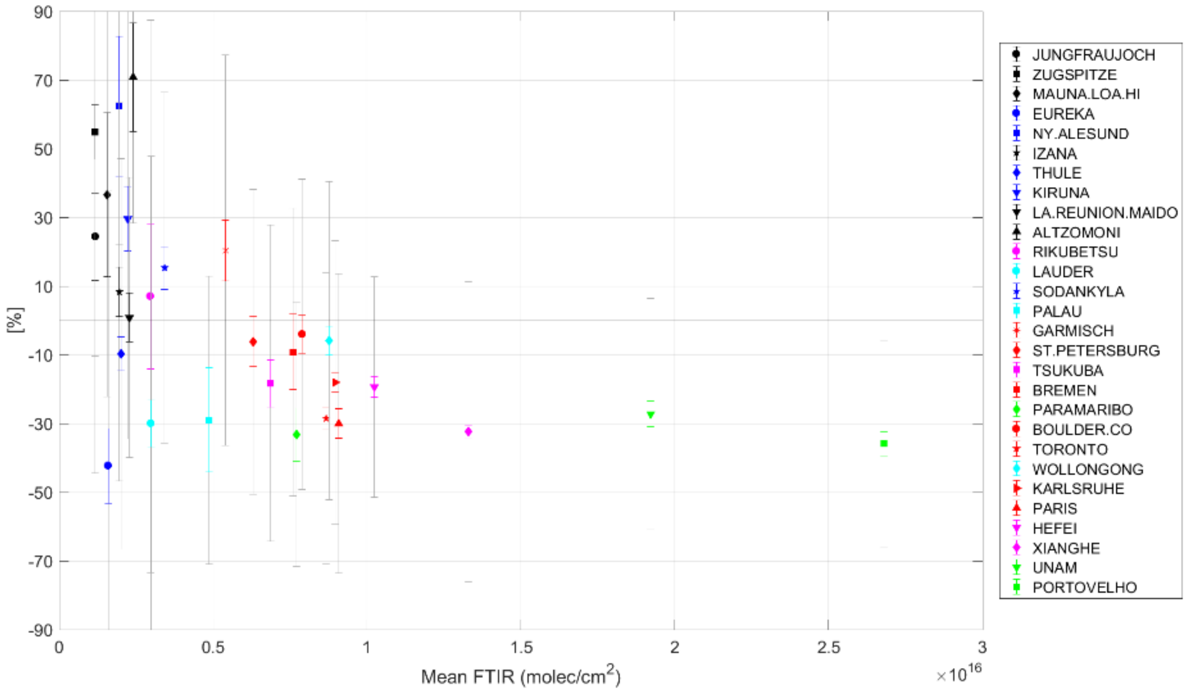


Figure 45: Bias at each station (in %) as a function of the mean FTIR total columns (10^{16} molec/cm^2). The gray bars are the systematic uncertainty on the differences, and the colored error bars are the 2-σ errors on the bias. The temporal range is May 2018 to December 2020 (updated from Vigouroux et al. (2020), Fig. 3).

7.3.3.2 MAX-DOAS UV-visible Spectrometers

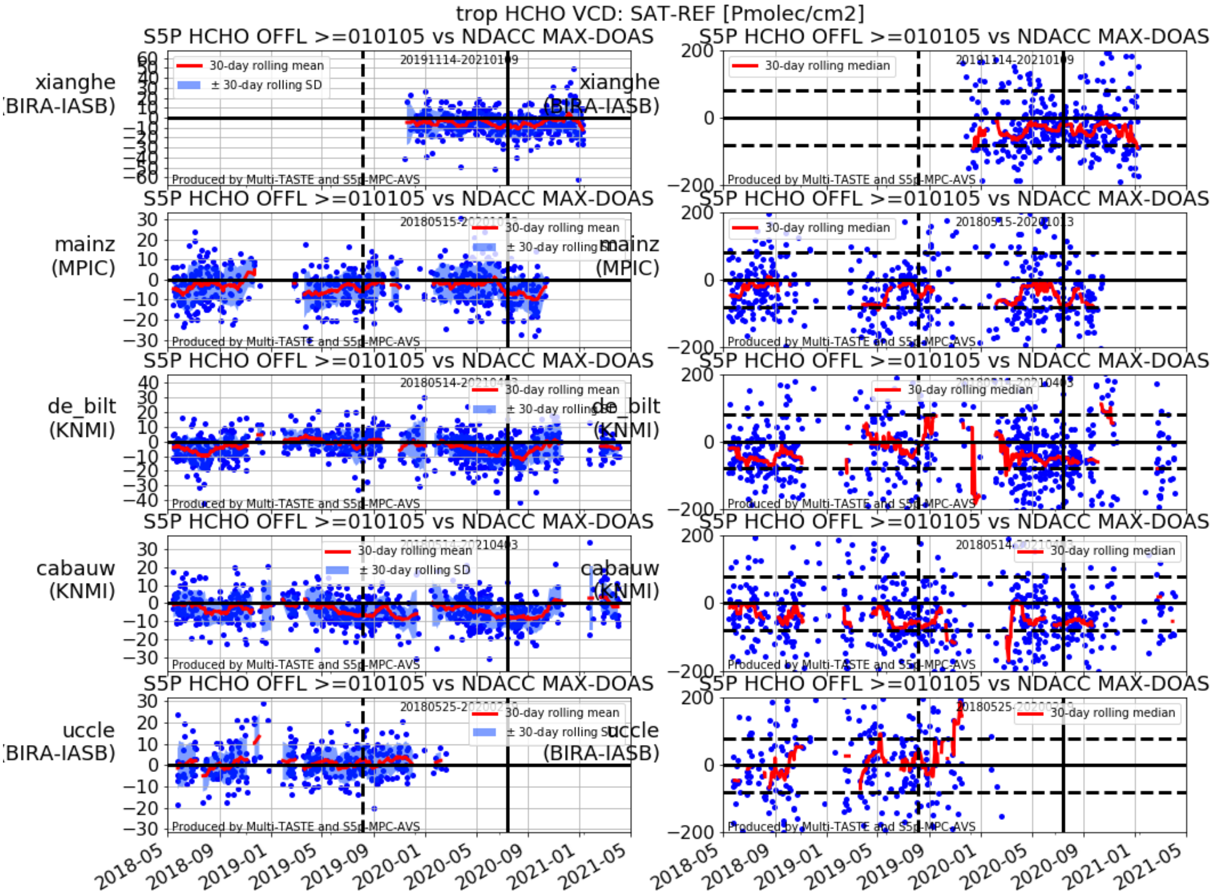


Figure 46: Time series of the difference [Pmolec/cm^2] and relative difference [%] between S5P RPRO+OFFL and MAX-DOAS HCHO column data at five stations. Data was obtained from the VDAF Automated Validation Server on 2021/05/27. Dashed line: pixel size switch at 2019/08/06. Full vertical line: OFFL processor version change on 2020/07/13.

Figure 45 shows difference time series S5P-MAXDOAS for the 5 sites from the AVS, with indication of the 30-day rolling monthly mean or median. Biases, if they occur, are mostly negative, but mostly within the 80% requirements. Pixel size switch and version change (v1 to v2, July 2020) have no impact on the comparison.

In the NIDFORVAL AO project, eight stations are used that provide data in the *GEOMS* format required in the project. As shown in **Figure 47** (upper panel), the median bias varies between -5% to -59%, with a median for all sites of -37%. Among these 8 stations, 3 are providing profiles and averaging kernels (AK) which allows 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 48**). For these 3 sites, the biases improve when using the averaging kernel (**Figure 47**, lower panel).

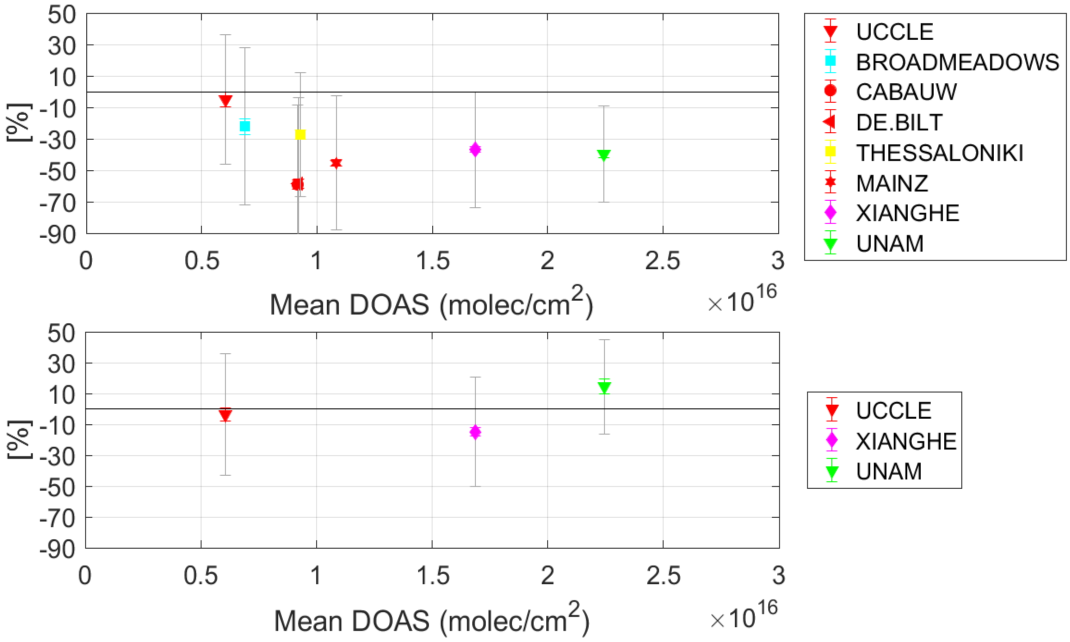


Figure 47: Bias at each station (in %) as a function of the mean DOAS total columns (10^{16} molec/cm^2). The grey bars are the systematic uncertainty on the differences, and the coloured error bars are the 2-σ error on the bias. Top panel: DOAS is the normal product without any modification. Bottom panel: DOAS data after Rodgers and Connor (2003) is applied (a priori substitution and smoothing with the TROPOMI averaging kernels).

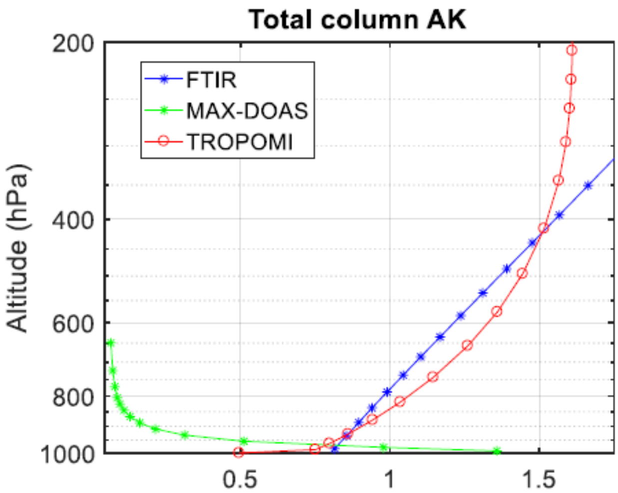


Figure 48: Typical total column averaging kernels for S5P TROPOMI and the two ground-based instrument types: FTIR (blue), MAX-DOAS (green), and TROPOMI (red). This illustrates the problem of the vertical smoothing difference error in these comparisons, as the instruments “see” different parts of the column.

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 DOAS smoothed comparisons +27±25% and -10±2%, respectively. The large uncertainty on the bias for low-HCHO levels is due to the small number of data involved, because the DOAS sites used in this study are not situated in a clean environment (see x-axis in **Figure 47**).

7.3.3.3 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.

7.3.3.4 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 49 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_v_clear).

When comparing N_v_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.

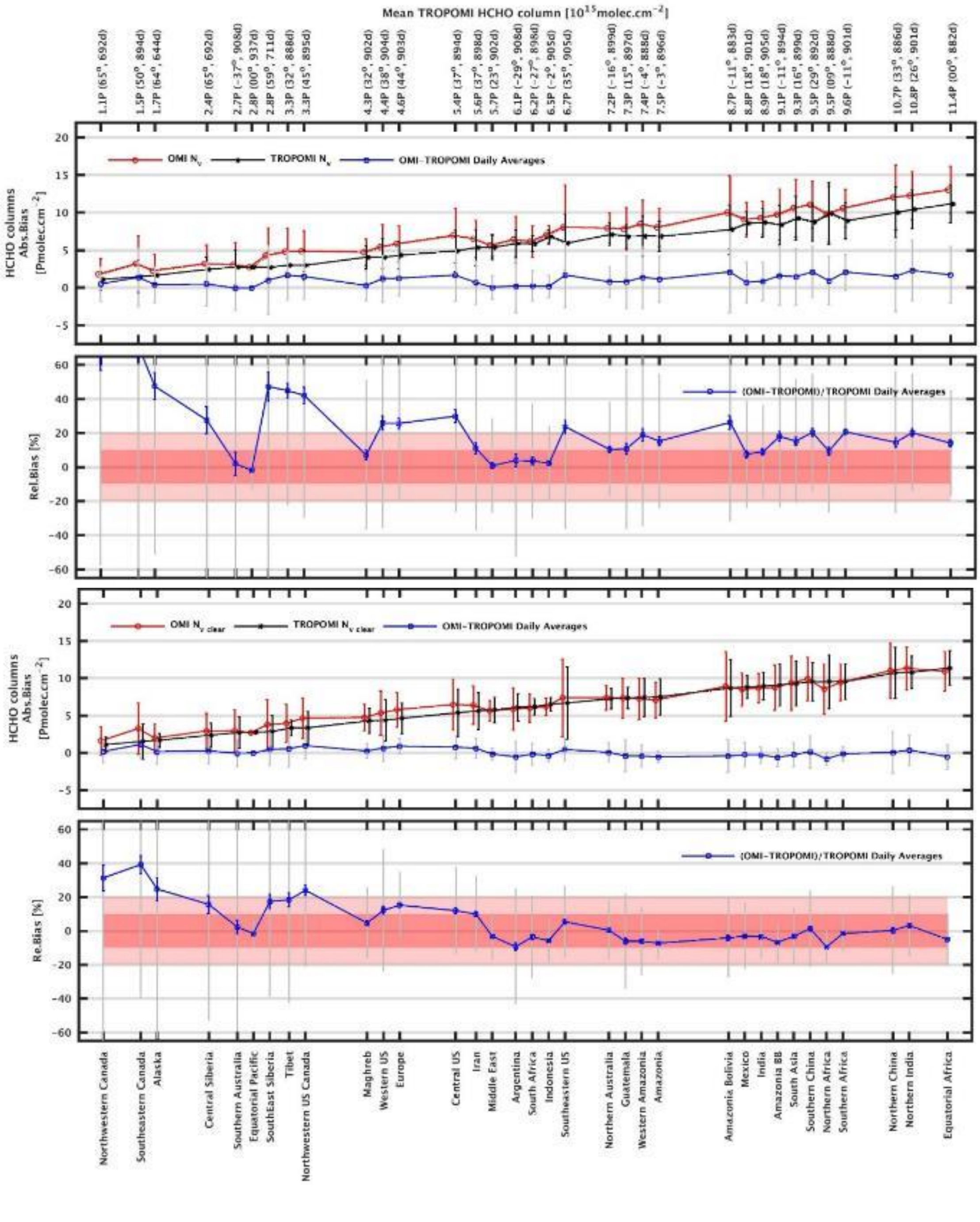


Figure 49: Absolute and relative biases between OMI and TROPOMI HCHO daily averaged tropospheric columns using cloud corrected AMF (N_v, two upper panels) or clear sky AMF (N_{v_clear}, two bottom panels) for the large regions represented on Figure 55. Regions are sorted as a function of the median TROPOMI HCHO column. Values of the averaged HCHO columns are provided on the top axis, as well as the numbers of common days taken for the comparison and the latitude of the region. The median OMI (red) and TROPOMI (black) columns are plotted together with the absolute differences (in blue). Error bars represent the median deviations of the columns, or the median absolute deviations of the differences (in grey). Pink areas indicate 10% and 20% bias.

7.3.4 Dispersion

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

7.3.4.1 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/cm². 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 2.42 Pmolec/cm², corresponding to a theoretical single-pixel dispersion of 2.4×√(#pixels)=12.9 Pmolec/cm², slightly above the 12 Pmolec/cm² requirement.

However, to evaluate the TROPOMI precision, it is more relevant to compare the MAD obtained at clean sites 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: 1.41 Pmolec/cm² for the 30-pixels-average, corresponding to a theoretical single pixel precision of 8 Pmolec/cm². 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 sites 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).

7.3.4.2 MAX-DOAS UV-visible Spectrometers

The MAD of the difference of S5P (single pixel) with respect to MAX-DOAS ranges from 7 Pmolec/cm² at Uccle to 10 Pmolec/cm² at Xianghe. This is within the mission requirement of precision of 12 Pmolec/cm². 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/cm² at Cabauw and De Bilt, respectively. This corresponds to a single pixel precision of 19-20 Pmolec/cm², which is twice larger than the pre-launch requirement of precision. However, if we look at the cleanest DOAS site Uccle to avoid larger collocation errors, the MAD is 2.6 Pmolec/cm² for the 34-pixels average comparisons, leading to a single pixel precision of 14 Pmolec/cm². If we apply the Rodgers and Connor (2003) technique, the MAD between TROPOMI and the $DOAS_{smoo}$ data is reduced at all the five sites, except at Uccle where the smoothing has little effect (**Figure 47**), leading to a single pixel precision of 15 Pmolec/cm² there.

In summary, the dispersion of the difference (7 to 10 Pmolec/cm², mainly polluted sites) is already within the dispersion requirement of 12 Pmolec/cm² for single-pixel comparisons with MAX-DOAS. Even lower dispersions (2 to 4 Pmolec/cm²) 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 √(#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 sites. The theoretical single pixel dispersion (~8 Pmolec/cm²) at clean FTIR sites is within the dispersion requirement, but 1.6 times larger than the random uncertainty budget provided in the TROPOMI files.

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

The mission requirement for S5P HCHO on dispersion is 12 Pmolec/cm². The number is valid at a single-pixel level. The single-pixel comparisons with MAX-DOAS from the VDAF validation server (mainly polluted sites) have a single-pixel dispersion of the difference of 8 to 10 Pmolec/cm². This is only an upper bound to the S5P 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/cm², confirming that at single-pixel level, the dispersion requirement is met.

Even lower dispersions (2 to 4 Pmolec/cm²) 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 √(#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 sites. It is therefore more appropriate to focus on clean FTIR sites. The theoretical single pixel dispersion (~8 Pmolec/cm²) at clean FTIR sites 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.

7.3.4.4 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/cm², respectively. When averaging data over large regions, the dispersion due to random uncertainties is greatly reduced compared to individual observations. As summarized in **Table 10**, the median absolute deviations of the monthly averaged columns are equivalent for OMI and TROPOMI (1.8 Pmolec/cm²), while the median deviations of their differences are significantly lower (0.5 Pmolec/cm²). 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/cm², while only 1 Pmolec/cm² 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 10 - Median absolute deviation of the OMI and TROPOMI daily and monthly averaged columns (N_{v_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.

Dispersion	OMI MAD [10^{15} molec.cm^{-2}]	TROPOMI MAD [10^{15} molec.cm^{-2}]	OMI-TROPOMI MAD [10^{15} molec.cm^{-2}]
Monthly Regional	1.8	1.8	0.5
Daily Regional	2.7	2.2	1.6
Monthly 20km	3.3	2.5	2.4
Daily 20km	7.8	3.7	7.1
Daily 20km in the Equatorial Pacific	3.5	1.0	3.7

7.3.5 Dependence on influence quantities

None to report.

7.3.6 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 50** 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 sites is 0.91.

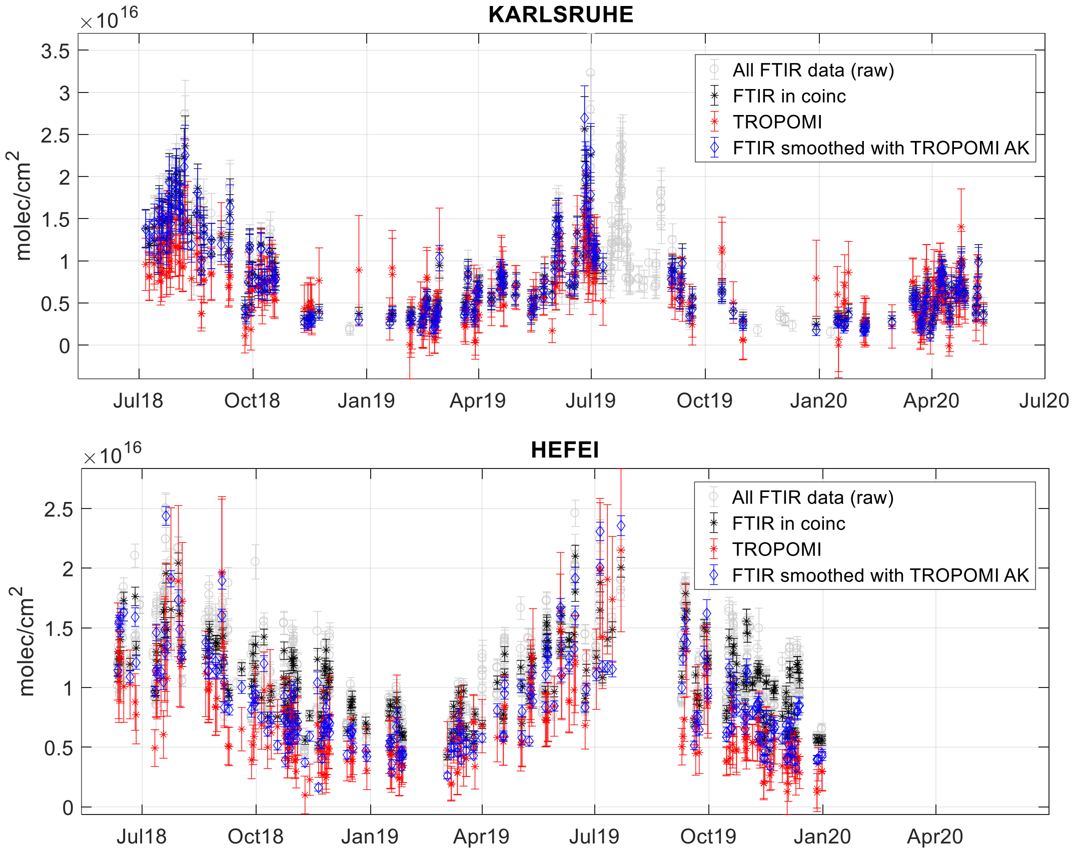


Figure 50: TROPOMI and FTIR HCHO time series at Karlsruhe and Hefei.

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 sites are shown in **Figure 51**.

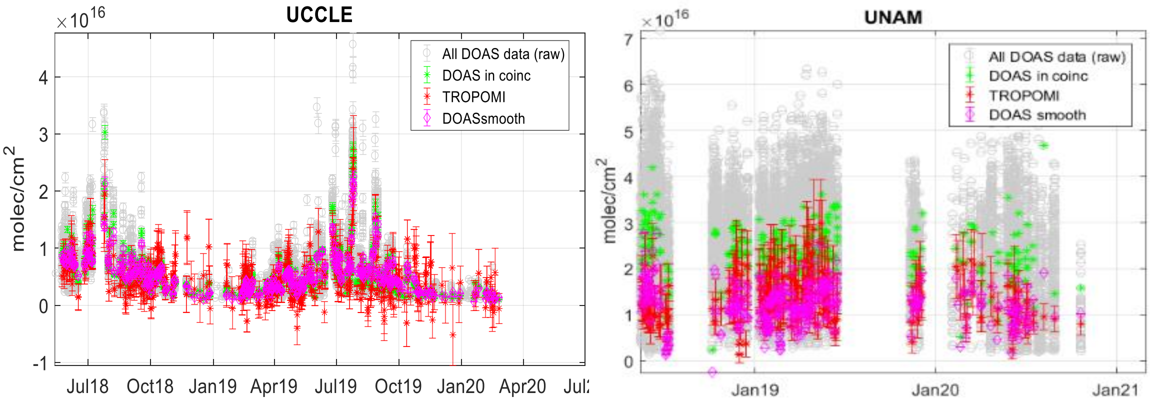


Figure 51: TROPOMI and MAX-DOAS HCHO time series at Uccle and Unam.

OMI QA4ECV comparisons

Day to day correlation between OMI and TROPOMI is very high above emission regions (**Figure 51)**.

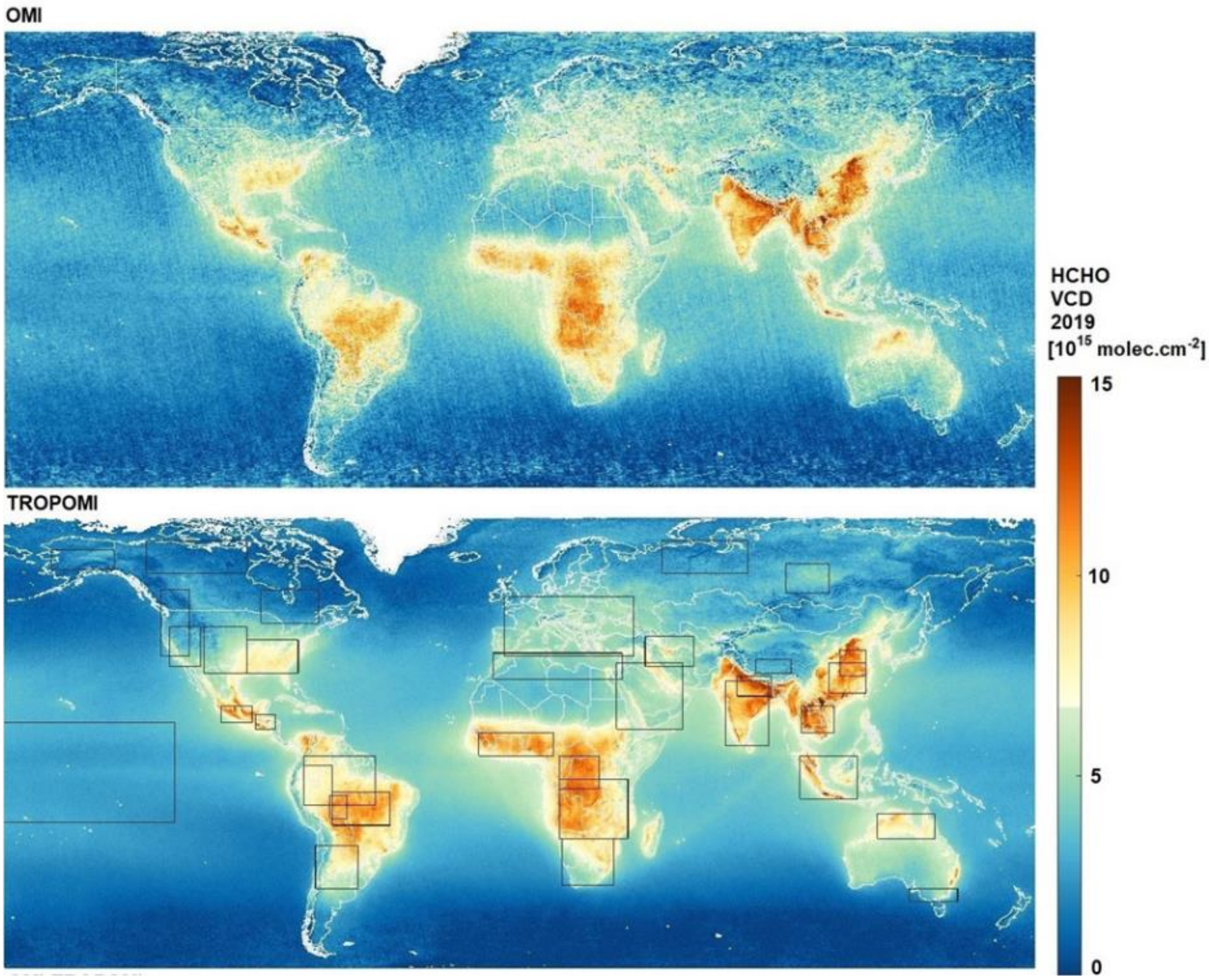


Figure 52: Average HCHO tropospheric column (N_{v_clear}) retrieved from OMI (first line) and TROPOMI (second line) in 2019. Limits of the regions selected for the comparisons are shown on the TROPOMI map. The same grid is used for both dataset (0.05°). Data are filtered using the product quality flags.

7.3.7 Geographical patterns

Figure 52 shows the comparison of OMI and TROPOMI HCHO columns (N_{v_clear}) averaged over one full year (2019). We observe a very good agreement overall. Differences range from 2 Pmolec/cm^2 over Tropics to -2 Pmolec/cm^2 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.

7.3.8 Impact of UPAS processor upgrade to version 02.01

The processor version changes to V02.01.03 in July 2020 and V02.01.04 in December 2020 does not impact the quality of the data as shown in the time evolution of the zonal mean HCHO daily columns (**Figure 53**). In the time series of HCHO differences for five MAX-DOAS stations (**Figure 54**) no impact of the version change is seen. The same conclusion is reached for comparisons with FTIR stations (not shown).

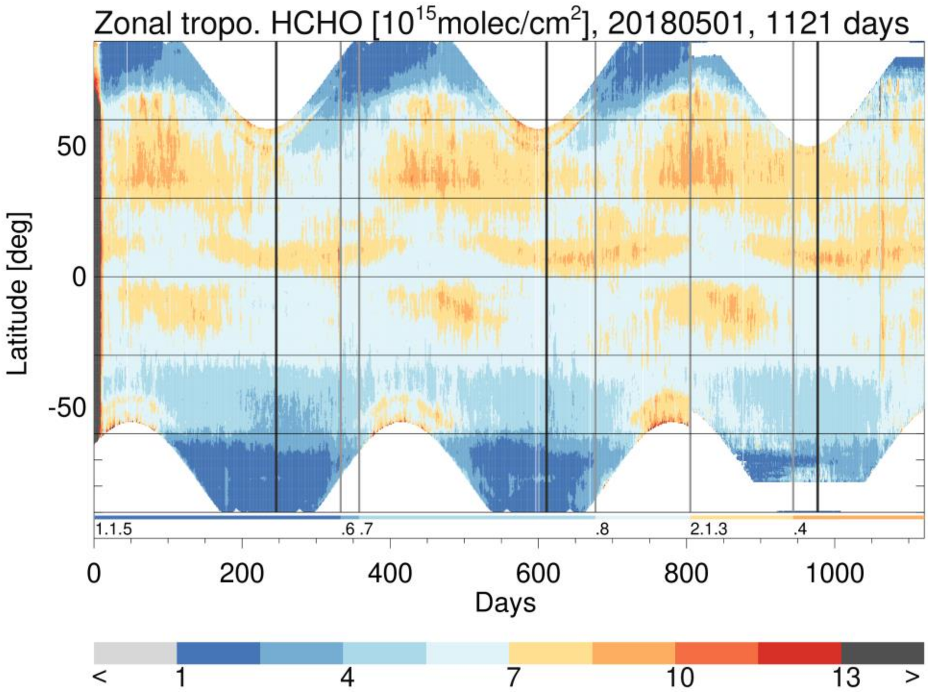


Figure 53: S5P TROPOMI HCHO columns [10^{15} molec/cm^2] as a function of day and latitude. The period is from May 2018 until mid-May 2021. Grid box size in latitude direction is 0.5°. The gray vertical lines mark the processor version changes, the black lines the beginning of each year.

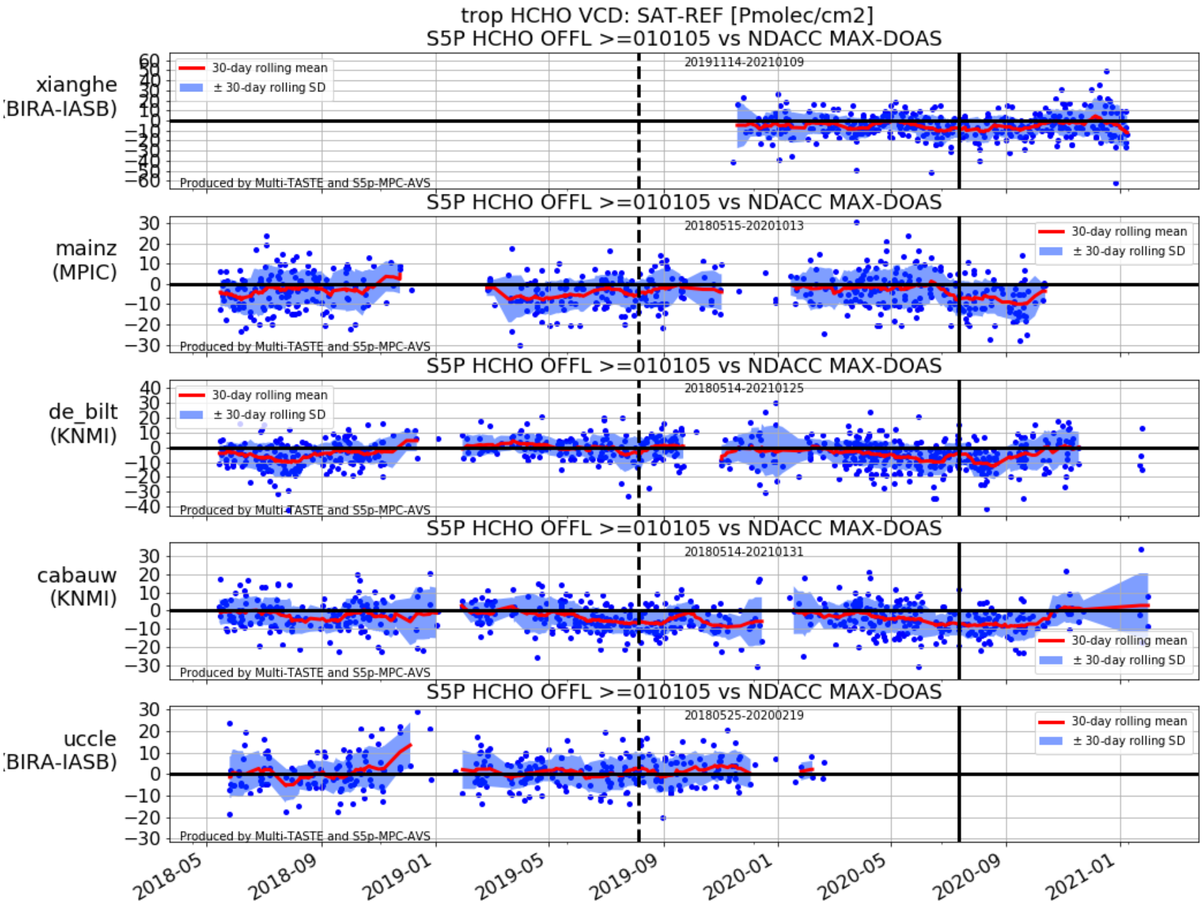


Figure 54: Time series of the difference [Pmolec/cm^2] between S5P RPRO+OFFL and MAX-DOAS HCHO column data at five stations. Data was obtained from the VDAF Automated Validation Server on 2021/02/23. Dashed line: pixel size switch at 2019/08/06. Full vertical line: OFFL processor version change on 2020/07/13.

7.4 Equivalence of L2_HCHO NRTI and OFFL products

We demonstrate the closeness of L2_HCHO NRTI and OFFL products at the MAX-DOAS sites De Bilt and Cabauw. NRTI (V01.01.02 to V02.01.04) and OFFL (RPRO V01.01.05 and OFFL V01.01.05 to V02.01.04) L2_HCHO results, each co-located with MAX-DOAS, were obtained from the VDAF Automated Validation Server. The subset of pixels, common to both NRTI and OFFL, was determined and differences between NRTI, OFFL, and MAX-DOAS determined. The statistical results are summarized in **Table 11.**

Table 11 - Statistics on the comparison of the common subset of L2_HCHO NRTI, L2_HCHO RPRO+OFFL and co-located MAX-DOAS for the sites Cabauw and De Bilt (*: unit of Pmolec cm^{-2}). The Automated Validation Server was consulted on 2021/02/23.

Cabauw: 510 common co-locations Orbits range from 4839 (2018-09-19) to 17112 (2021-01-31)			
	NRTI vs OFFL	NRTI vs MXD	OFFL vs MXD
Mean(diff)±sem*	-0.08	-4.12±0.37	-4.04±0.37
Median(diff)*	-0.02	-4.76	-4.35
Std(diff)*	1.6	8.4	8.4
1/2 IP68(diff)*	1.2	7.5	7.6
Pearson R	0.98	0.35	0.35
Slope	0.98	0.51	0.51

De Bilt: 488 common co-locations Orbits range from 4839 (2018-09-19) to 17026 (2021-01-25)			
	NRTI vs OFFL	NRTI vs MXD	OFFL vs MXD
Mean(diff)±sem*	-0.11	-2.64±0.42	-2.53±0.43
Median(diff)*	-0.00	-2.62	-2.30
Std(diff)*	2.2	9.3	9.4
1/2 IP68(diff)*	1.4	9.0	8.6
Pearson R	0.97	0.25	0.24
Slope	0.96	0.35	0.35

7.4.1 Bias

At the MAX-DOAS sites, 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 11**). 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 difference in bias between NRTI and OFFL is therefore not statistically significant. The same conclusions are found using the FTIR network (Vigouroux et al. 2020). An example of the spatial distribution of daily differences is shown in **Figure 55** (2021/05/15) for the newest V02.01.04. The relative difference is -1.3% with a Pearson correlation coefficient R of 0.97.

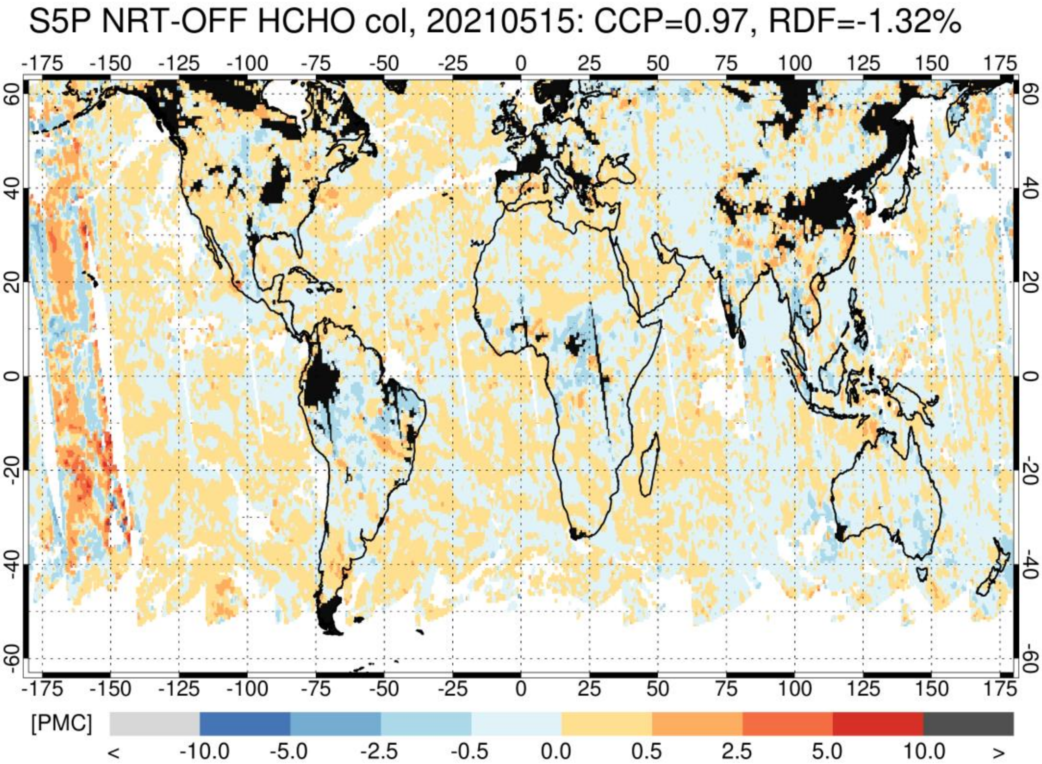


Figure 55: Difference between S5P V02.01.04 NRTI and OFFL HCHO column data for 2021/02/01 (binned to 0.5°x0.5° resolution). The difference is only calculated for columns above 0.5 Pmolec/cm^2.

7.4.2 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.

8 Validation Results: L2_SO2

8.1 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 8.4 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 version 02.01.04 released in November 2020 show good consistency of the SO_2 data product before and after the processor switch, with rather modest effects.

8.2 Validation approach

8.2.1 Ground-based networks

Boundary layer pollution (SO_2 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 SO_2 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. SO_2 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 SO_2 concentration gradients in volcanic plumes, the comparison is not performed using the SO_2 columns but rather using the derived SO_2 fluxes.

8.2.2 Satellites

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

8.2.3 Field campaigns and modelling support

S5P TROPOMI L2_SO2 sulphur dioxide column data are compared to car MAX-DOAS measurements performed in Lahore.

8.2.4 Test of the expectation of zero SO_2 SCDs (within detection limit) outside volcanic plumes and strongly polluted regions

Outside strongly polluted regions and volcanic plumes, the atmospheric SO_2 concentrations are very low and the corresponding SO_2 columns are below the detection limit of S5P TROPOMI. Thus S5P TROPOMI measurements outside strongly polluted regions and volcanic plumes are used to check the consistency of the S5P TROPOMI L2_SO2 sulphur dioxide column data with the assumption of SO_2 slant column densities (SCD) of zero. From this test, also the spread of the S5P TROPOMI L2_SO2 sulphur dioxide column data is quantified.

8.3 Validation of L2_SO2 NRTI

8.3.1 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_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/products-algorithms

8.3.2 Status of validation

So far the validation of the S5P TROPOMI L2_SO2 sulphur dioxide column data is mainly based on satellite to satellite comparisons (**Figure 56**, **Figure 57**), 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 58** to **Figure 63**) 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_2 columns were mostly close to or below the detection limit of S5P TROPOMI.

S5P TROPOMI L2_SO2 sulphur dioxide column data were also compared to ground based MAX-DOAS measurements from the NOVAC network. However, the SO_2 columns were not compared directly, because of the strong gradients across volcanic plumes. Instead the derived SO_2 fluxes were compared, for which good agreement was found.

Outside strongly polluted regions and volcanic plumes, the atmospheric SO_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.

Finally, first verification checks of the UPAS processor upgrade to version 02.01.03 released on 16 July 2020 (and subsequent version 02.01.04) 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 64** and **Figure 66**).

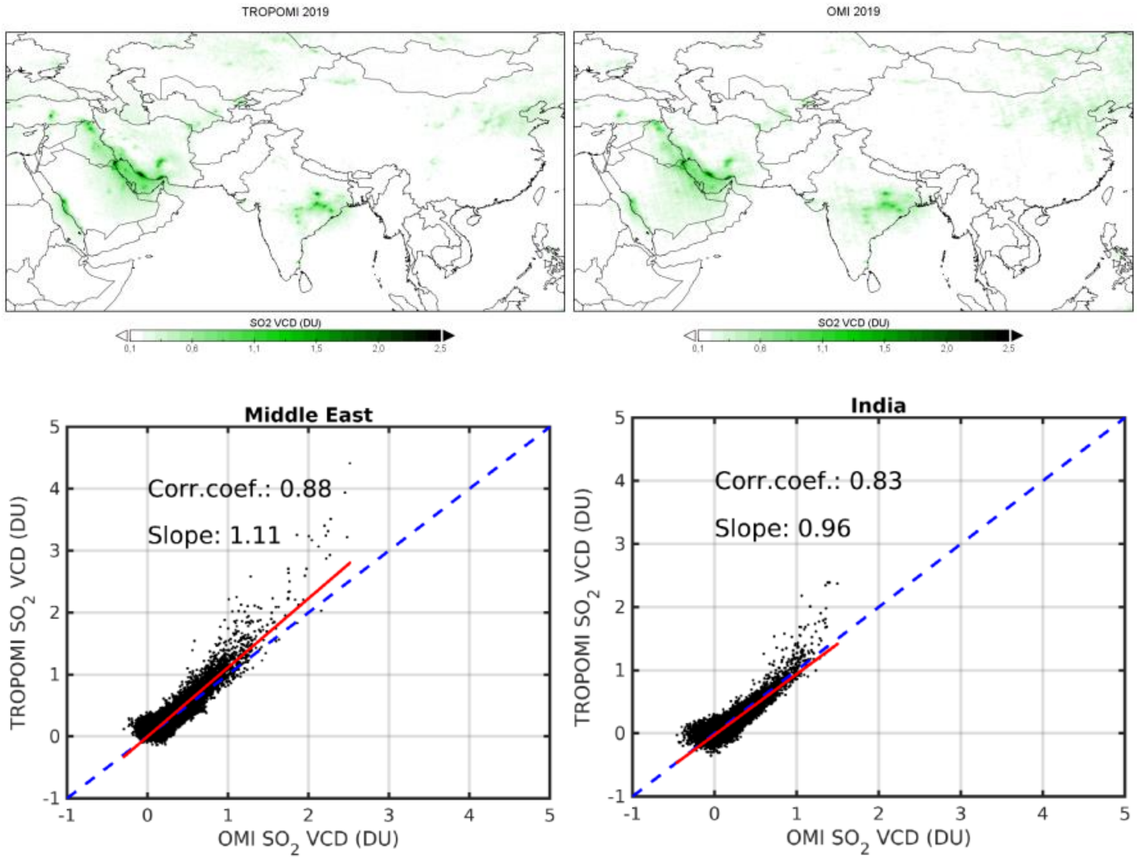


Figure 56: Top: Comparison of the average distribution (01 Jan 2019 – 31 Dec 2019) of the SO_2 VCDs derived from TROPOMI and OMI over regions with strong air pollution. Both data sets show very good agreement. Bottom:Correlation plots TROPOMI versus OMI over the Middle East and India. Note that a fixed AMF of 0.4 was used for both retrievals to exclude the effect of different profile assumptions. Courtesy of Nicolas Theys, BIRA-IASB.

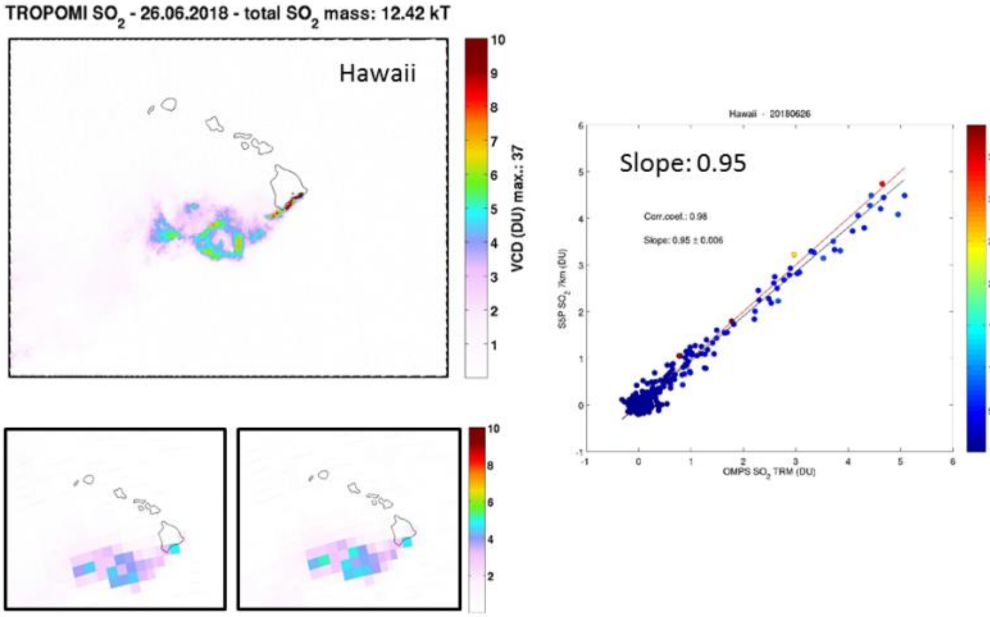


Figure 57: Comparison of TROPOMI and OMPS measurements of the volcanic plume of Kilauea on 26 June 2018. The large figure shows the original TROPOMI data. The two small figures show the spatially degraded TROPOMI data and the OMPS data. The figure right shows the correlation plot of the degraded TROPOMI data versus the collocated OMPS data. Courtesy of C. Li and N. Krotkov, NASA/GSFC.

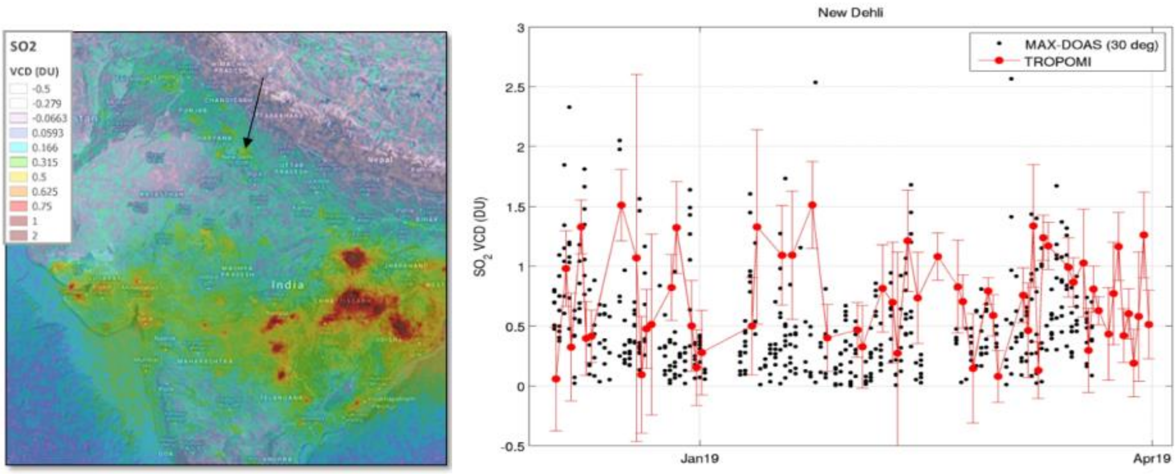


Figure 58: Comparison of TROPOMI SO_2 VCDs to MAX-DOAS measurements at Greater Noida (close to New Delhi, India). The following selection criteria were applied: distance < 15km, CF<0.2, AMF>0.2, MAX-DOAS +/- 1h around S5P overpass. Courtesy of M. Sharma (Sharda University, India) S. Donner, S. Dörner, T. Wagner (MPIC), N. Theys (BIRA-IASB).

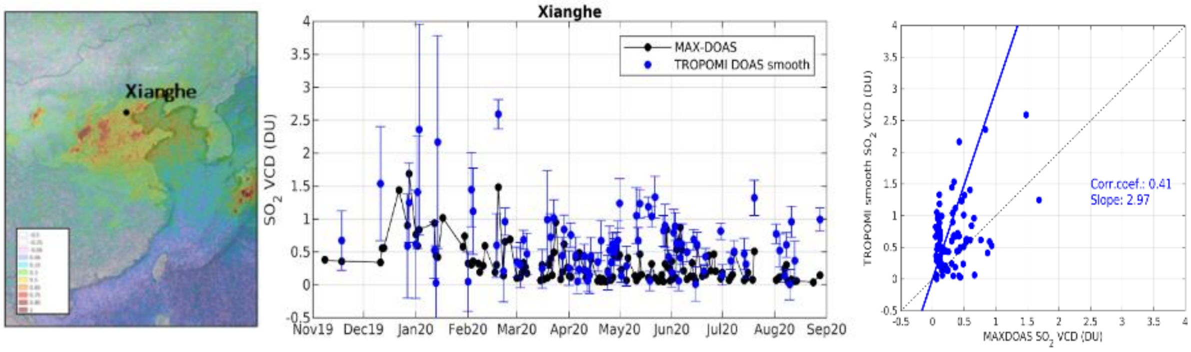


Figure 59: Comparison of TROPOMI SO_2 VCDs to MAX-DOAS measurements (daily means) at Xianghe (China). The following selection criteria were applied: distance < 25km, CF<0.2, AMF>0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB).

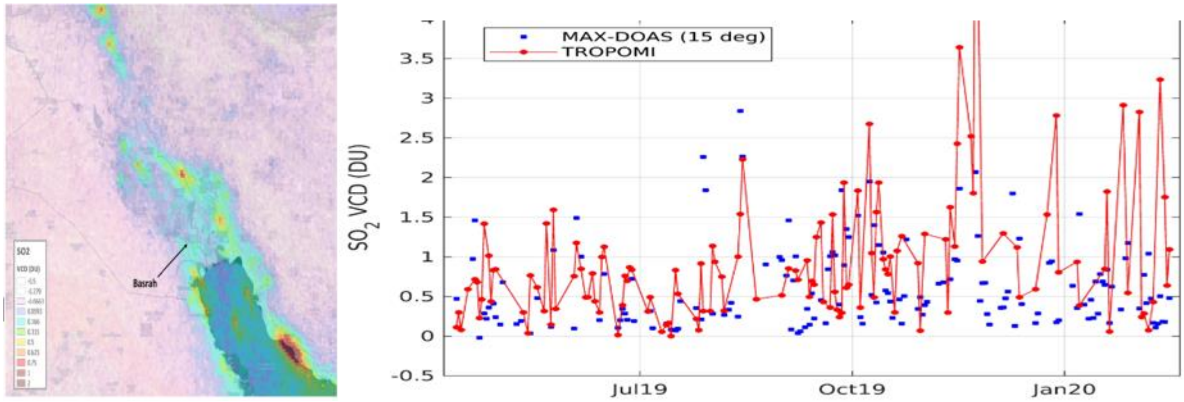


Figure 60: Comparison of TROPOMI SO_2 VCDs to MAX-DOAS measurements (daily means) at Basra (Iraq) from April 2019 to March 2020. The following selection criteria were applied: distance < 25km, CF<0.2, AMF>0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB), data provided by Nayyef Almaliki, Mustafa Aldossary, Ali Almasoudii, Sebastian Donner, Steffen Dörner, Thomas Wagner.

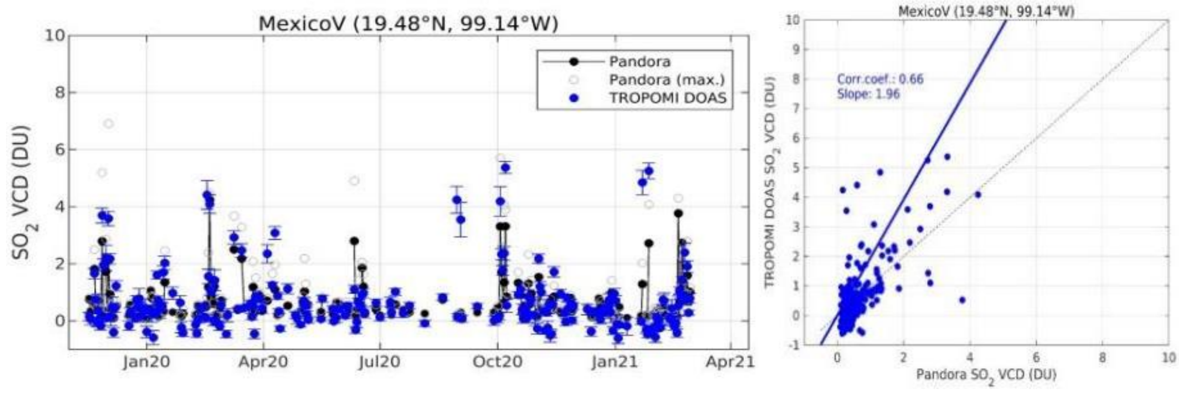


Figure 61: Comparison of TROPOMI SO_2 VCDs to preliminary Pandora measurements (daily means) at Mexico City from September 2019 to March 2021. The following selection criteria were applied: distance < 25km, CF<0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB), Martin Tiefengraber and Alexander Cede (Luftblick). Pandora data are still in demonstration phase.

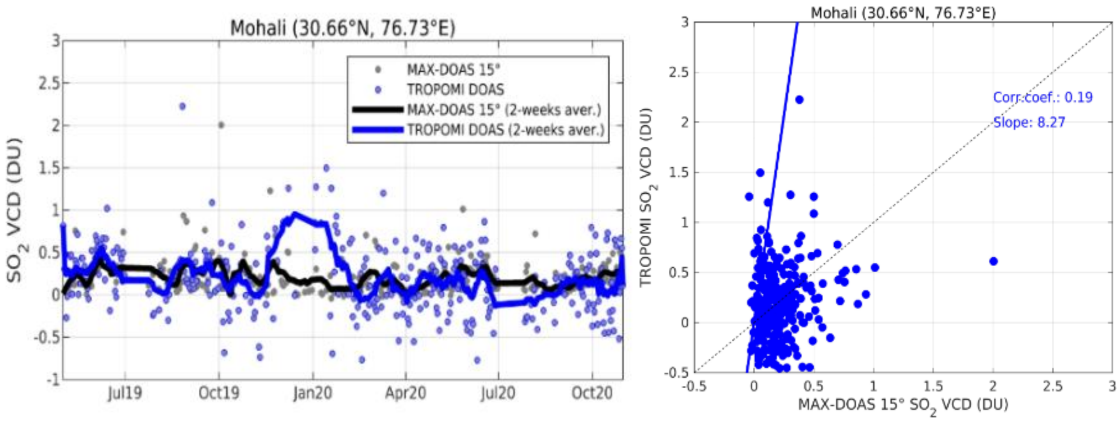


Figure 62: Comparison of TROPOMI SO_2 VCDs to MAX-DOAS measurements (daily means) at Mohali (India) from June 2019 to November 2020. The following selection criteria were applied: distance < 25km, CF<0.2, AMF>0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB), data provided by, Vinod Kumar, Vinayak Sinha, Sebastian Donner, Steffen Dörner, Thomas Wagner.

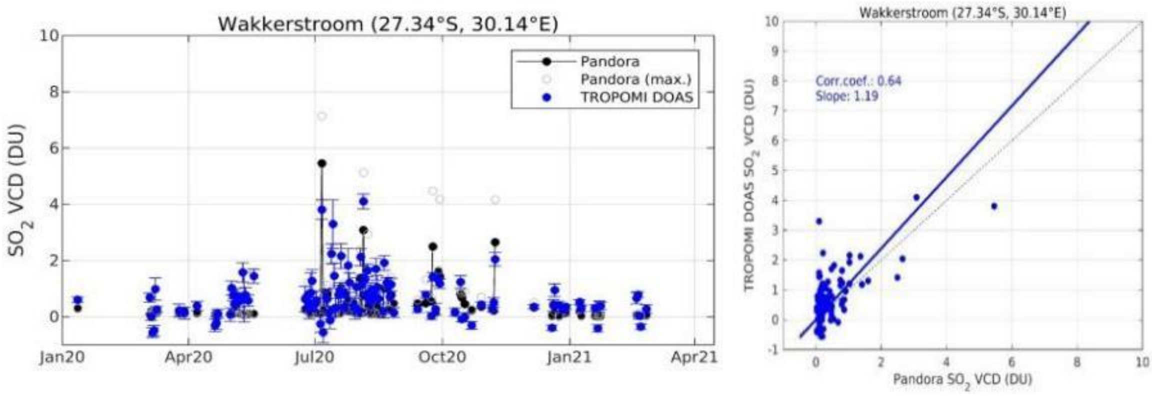


Figure 63: Comparison of TROPOMI SO_2 VCDs to preliminary Pandora measurements (daily means) at Wakkerstroom (South Afrika) from March 2020 to March 2021. The following selection criteria were applied: distance < 25km, CF<0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB), Martin Tiefengraber and Alexander Cede (Luftblick). Pandora data are still in demonstration phase.

8.3.3 Bias

The bias is well within requirements for observations of volcanic plumes and boundary pollution. From the time series of averaged SO_2 SCDs it is estimated that the bias is within 0.2 DU.

8.3.4 Dispersion

The dispersion is well within requirements for observations boundary pollution. For observations of strong volcanic plumes the dispersion can be slightly above the requirements. The slightly larger dispersion over strong volcanic plumes is not seen as a substantial restriction of the data quality. From the time series of the standard deviation of the SO_2 SCDs it is estimated that the dispersion is within 0.2 DU.

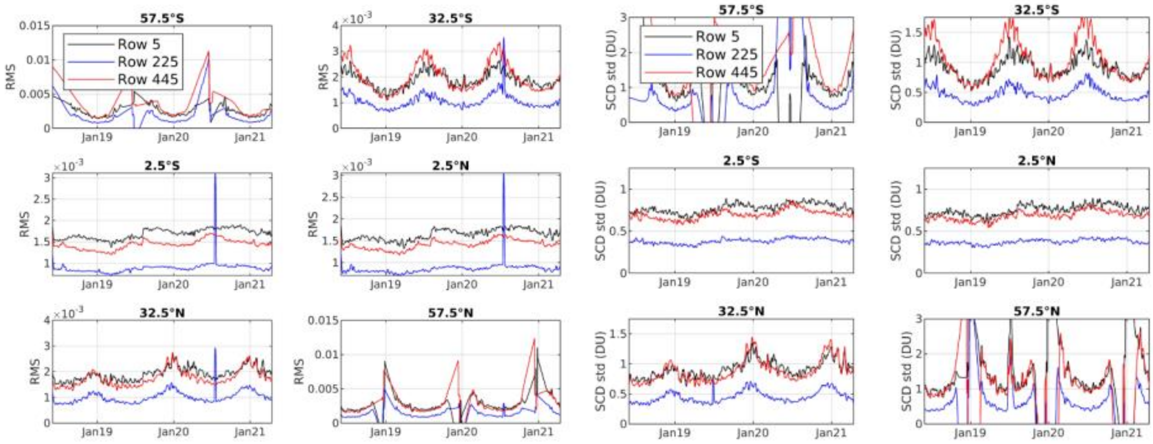


Figure 64: Temporal evolution of the measurement error (left) and the standard deviation (right) for selected 5° latitude bands and 3 detector rows from December 2018 to April 2021. Good qualitative agreement between both quantities is found indicating that the random uncertainty is well characterized 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).

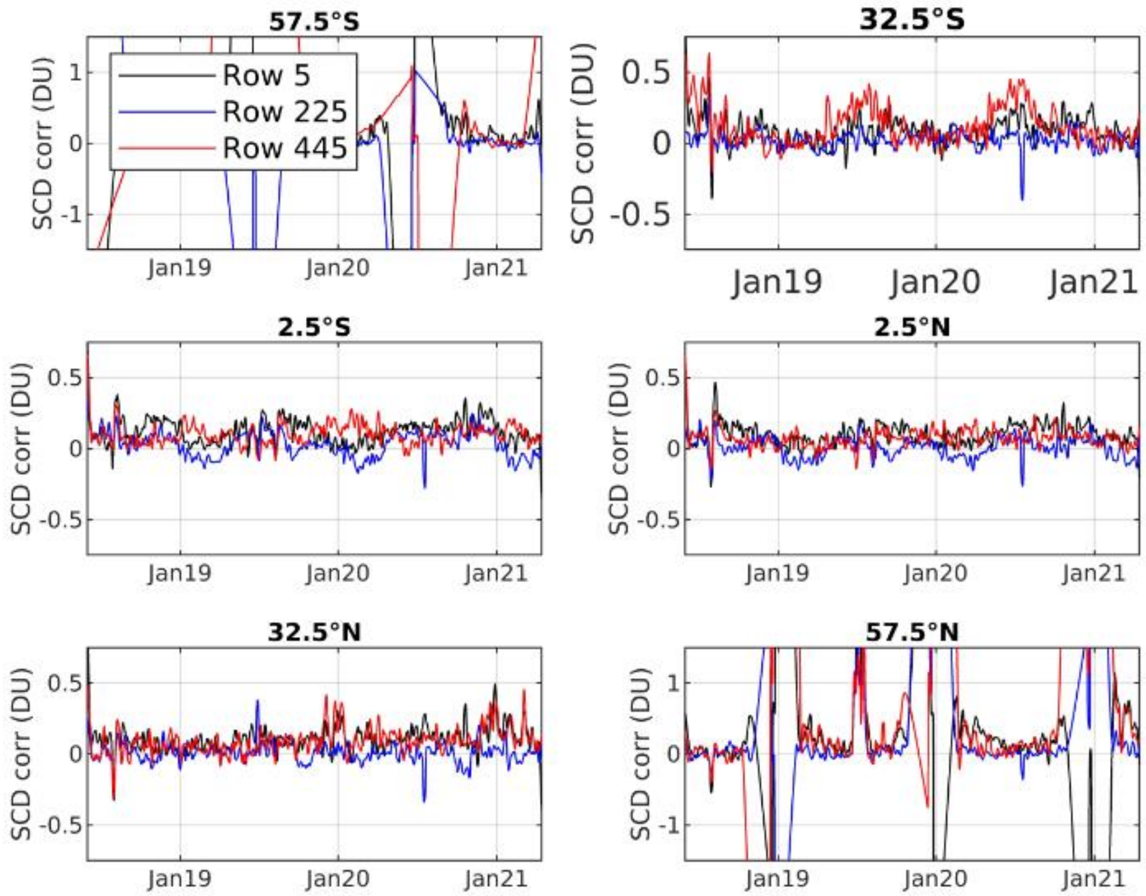


Figure 65: Temporal evolution of the averaged SO_2 SCD for selected 5° latitude bands and 3 detector rows from December 2018 to April 2021. The values are close to zero and show relatively small day to day variations. The larger variations in August 2019 are caused by strong volcanic eruptions. Courtesy of Nicolas Theys (BIRA-IASB).

8.3.5 Dependence on influence quantities

Slightly larger bias and dispersion are found towards higher SZA.

8.3.6 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.

8.3.7 Geographical patterns

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

8.3.8 Other features

None to report.

8.4 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.

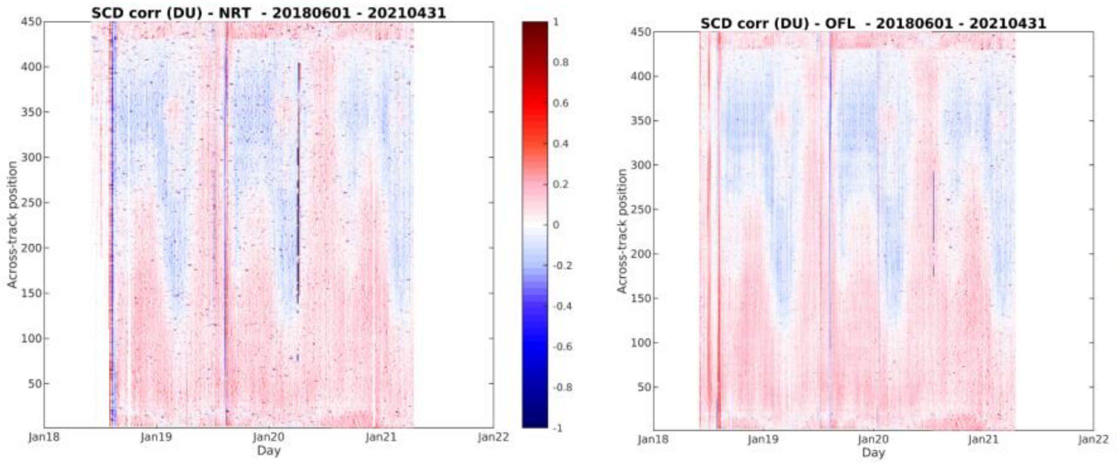


Figure 66: Comparison of the NRT (left) and offline (right) SO_2 data products. Shown are the time series of background corrected SO_2 SCDs for all 450 detector rows from June 2018 to April 2021. Most of the short time features (vertical colored lines) are caused by individual strong volcanic eruptions. 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. Courtesy of Nicolas Theys, BIRA-IASB.

.

9 Validation Results: L2_CO

9.1 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 9.4.

9.2 Validation approach

9.2.1 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) and the Total Carbon Column Observing Network (TCCON, https://tccondata.org). **Figure 67** displays the geographical distribution of the NDACC and TCCON sites. Near-infrared TCCON measurements provide CO column averaged (xCO) data with typical uncertainty values of 2% for the bias and 1% for the precision. Solar infrared NDACC measurements provide CO total column data with a typical total uncertainty of 3%.

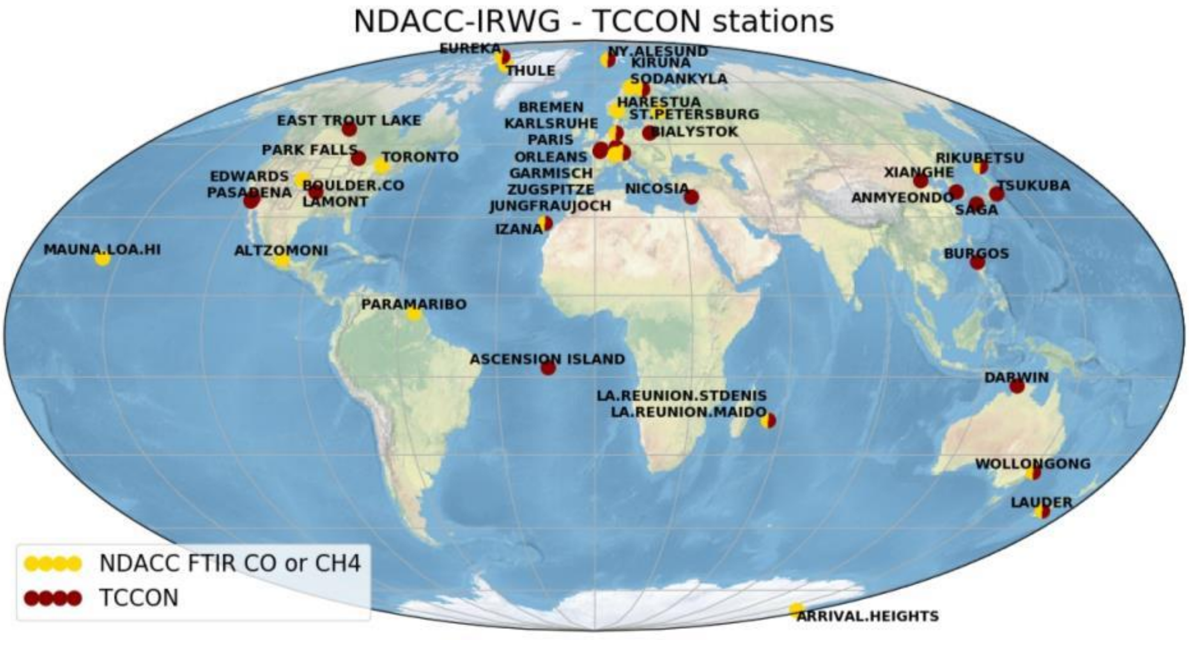


Figure 67: Geographical distribution of NDACC and TCCON FTIR stations measuring atmospheric carbon monoxide column data. Some sites contribute to both networks.

9.2.2 Satellites

None for this report.

9.2.3 Field campaigns and modelling support

None for this report.

9.3 Validation of L2_CO OFFL

9.3.1 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 data product, all available on https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms.

9.3.2 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 the validation methodology reported at the third 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.

Current conclusions are based on the amount of reference measurements available at the time of this analysis, yielding comparison pairs from November 2017 through March 2021. 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 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 3h for NDACC observations. The 1-hour interval can be justified by noting that TCCON instruments acquire only one type of spectra, while NDACC instruments are required to measure different types of spectra, making the number of CO observations more sparse. In the comparison procedure with TCCON data, the apriori in the TCCON retrievals have been substituted with the S5P CO apriori (Rodgers 2003). The validation procedure for both the NDACC and TCCON based comparisons includes an adaptation of the TROPOMI CO column to the altitude of the groundbased 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.

9.3.3 Bias

The systematic difference between S5P L2_CO daily mean data and correlative ground-based measurements is on an average 6.75% with respect to NDACC data and 9.14% with respect to TCCON data. At some sites like mountain stations those values are exceeded likely because of geographical colocation issues. This bias estimate falls well within the mission requirements. **Figure 69** and **Figure 70** show the biases over the time period Nov 2017– March 2021 sorted by latitude. It seems that there is no latitudinal dependence of the bias. **Figure 68** does not show any significant degradation in bias with time (note that the longer time period covers different processor versions). **Figure 68** also shows a slight increase of bias during local winter.

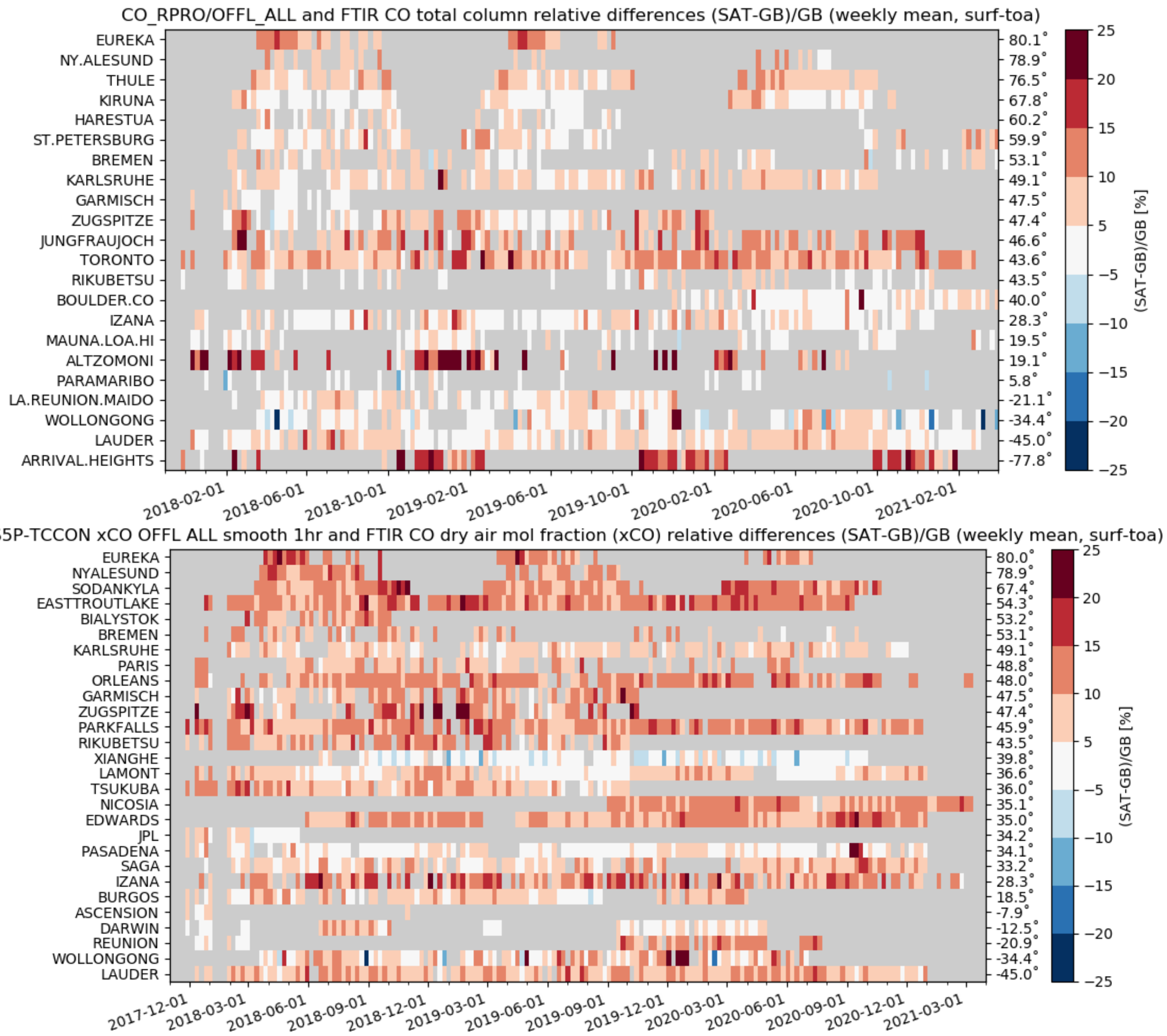


Figure 68: Relative bias between S5P L2_CO OFFL and ground-based CO column data at NDACC (top) and TCCON (bottom) FTIR stations. Over the Nov. 2017 – March 2021 time period the plots do not show a clear meridian dependence or temporal change in the weekly averaged biases.

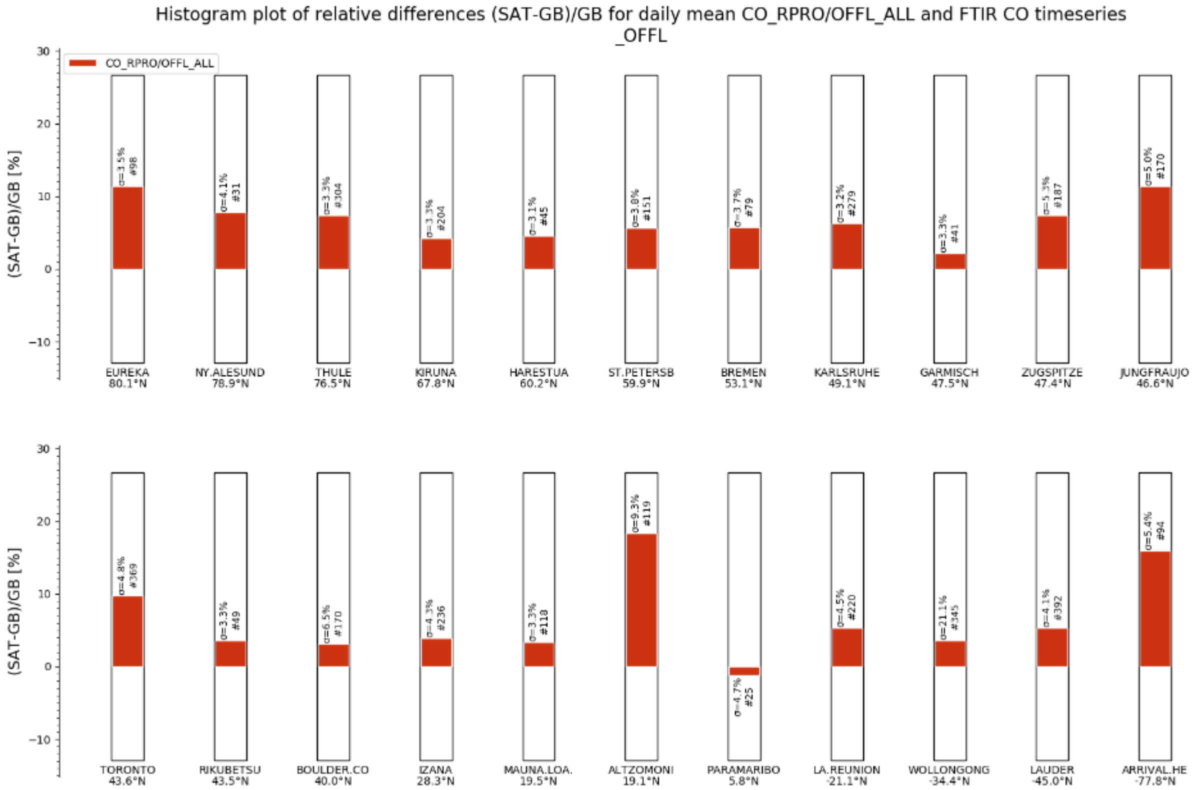


Figure 69: Bar chart of the relative mean difference between S5P L2_CO OFFL and FTIR CO column data at 22 NDACC sites, for all data within the time range from November 2017 till March 2021 showing RPRO/OFFL. The sites are sorted with decreasing latitude. All biases are below 15% except at Altzomoni, which is a mountain city near Mexico City: the higher bias is due to the chosen pixel selection criteria, here higher concentration pixels near the city are taken into account in the average. Arrival Height (Antarctic) also shows a larger bias.

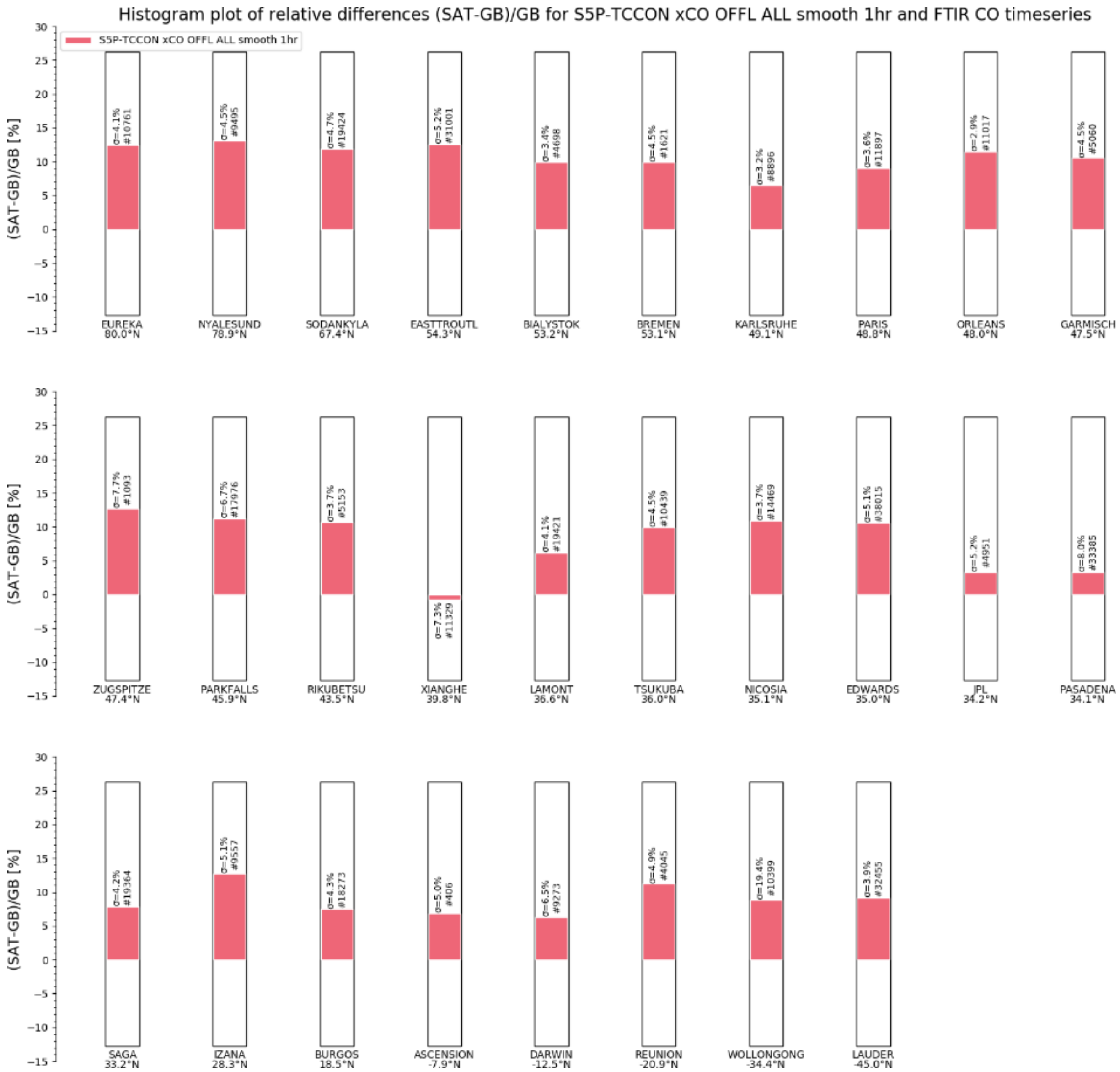


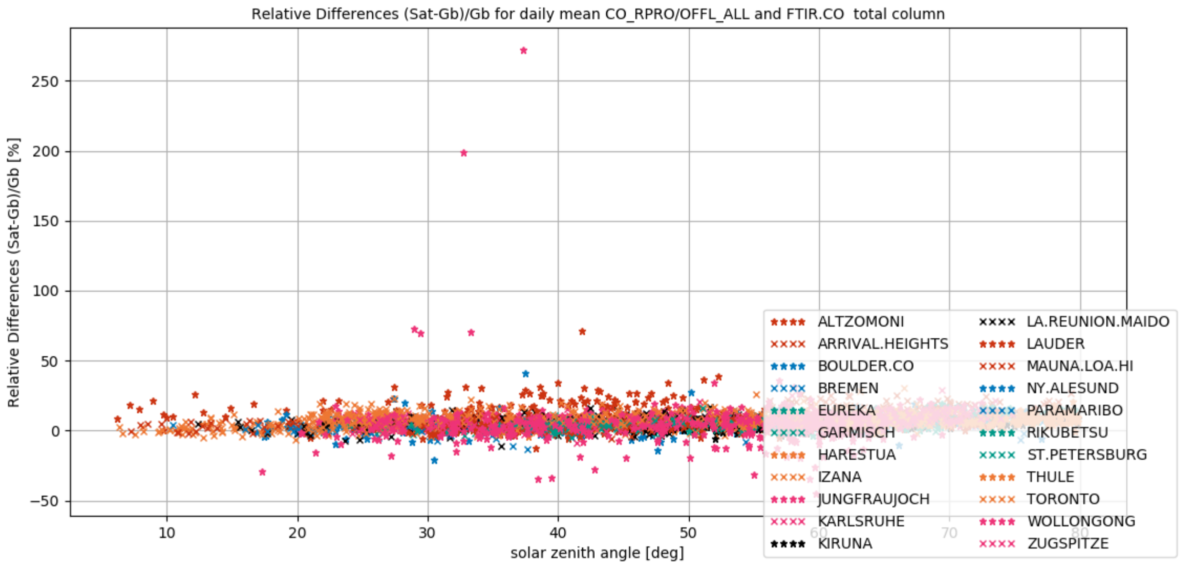
Figure 70: Bar chart of relative mean difference between S5P L2_CO OFFL and FTIR CO column data at 28 TCCON sites for all data within the time range Nov 2017 till March 2021. The sites are sorted with decreasing latitude. The majority of the OFFL biases are below 12% except in the Arctic and mountain station (Izaña and Zugspitze) where the bias is slightly above 12%. Xianghe station lies in a polluted region where we see almost zero bias.

9.3.4 Dispersion

The 1σ dispersion of the relative mean bias around its mean is of the order of 5%. The individual values for the different sites are indicated in **Figure 69** and **Figure 70.** This dispersion can be considered as an upper boundary of the random uncertainty of the satellite data.

9.3.5 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.



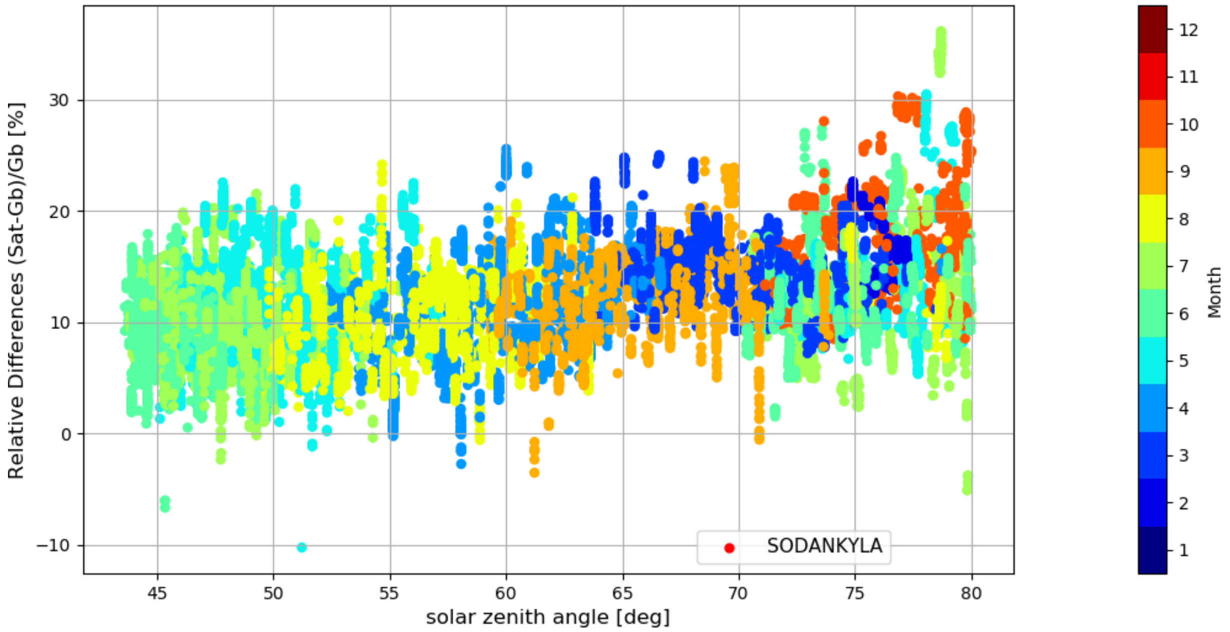
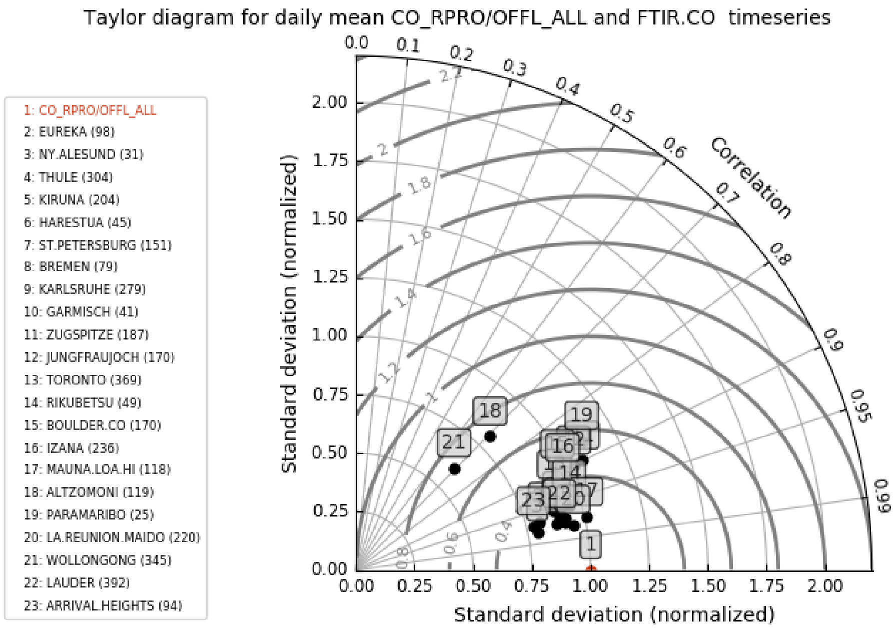


Figure 71: Relative difference (daily mean) between S5P L2_CO RPRO/OFFL and NDACC (top) or TCCON Sodankylä (bottom) carbon monoxide total column as a function of the TROPOMI solar zenith angle, in the 'all' case. The Wollongong time series contains some outlying values, which can be attributed to co-location mismatches during the Australian fires in Jan 2020.

9.3.6 Short term variability

For all the NDACC and TCCON 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 and on average reaching 0.9 (**Figure 72**).



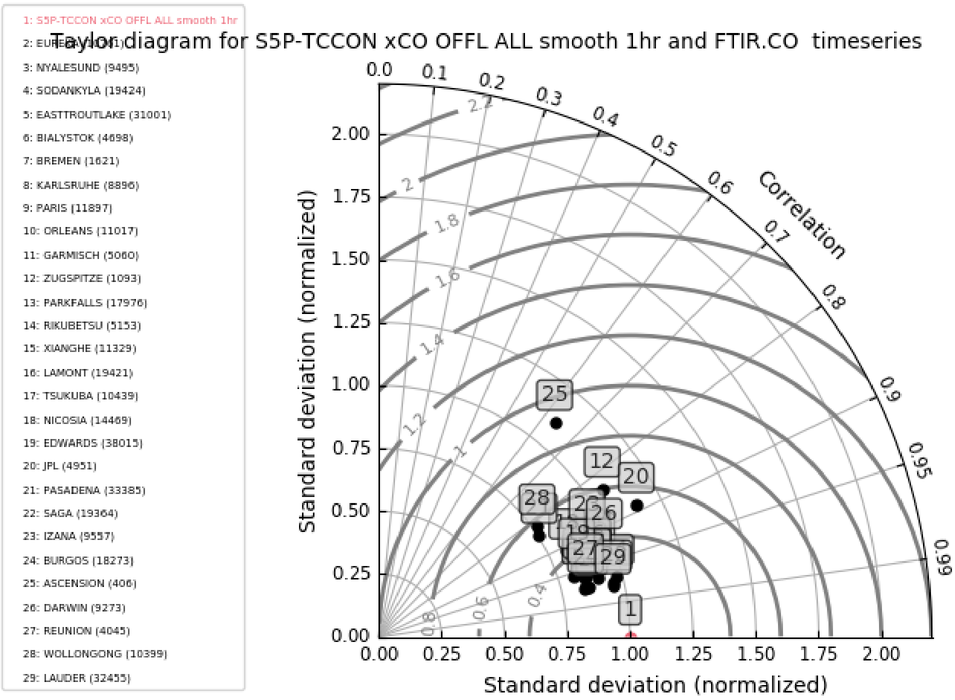


Figure 72: Taylor diagrams for daily mean differences between S5P L2_CO OFFL and ground-based networks CO data: NDACC (top) and TCCON (bottom) for our all case of pixel selection criteria (Boulder and Ascension cover a limited time period), NDACC Wollongong has reduced correlation due to co-location mismatches during the Australian fires in Jan 2020).

9.3.7 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 73** 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 also at preliminary validation results. How this can be removed from the data is discussed in the PRF and is subject to further investigation in the framework of instrument calibration.

TROPOMI CO column data also suffer from instrumental effects of the South Atlantic Anomaly (SAA), see **Figure 74**.

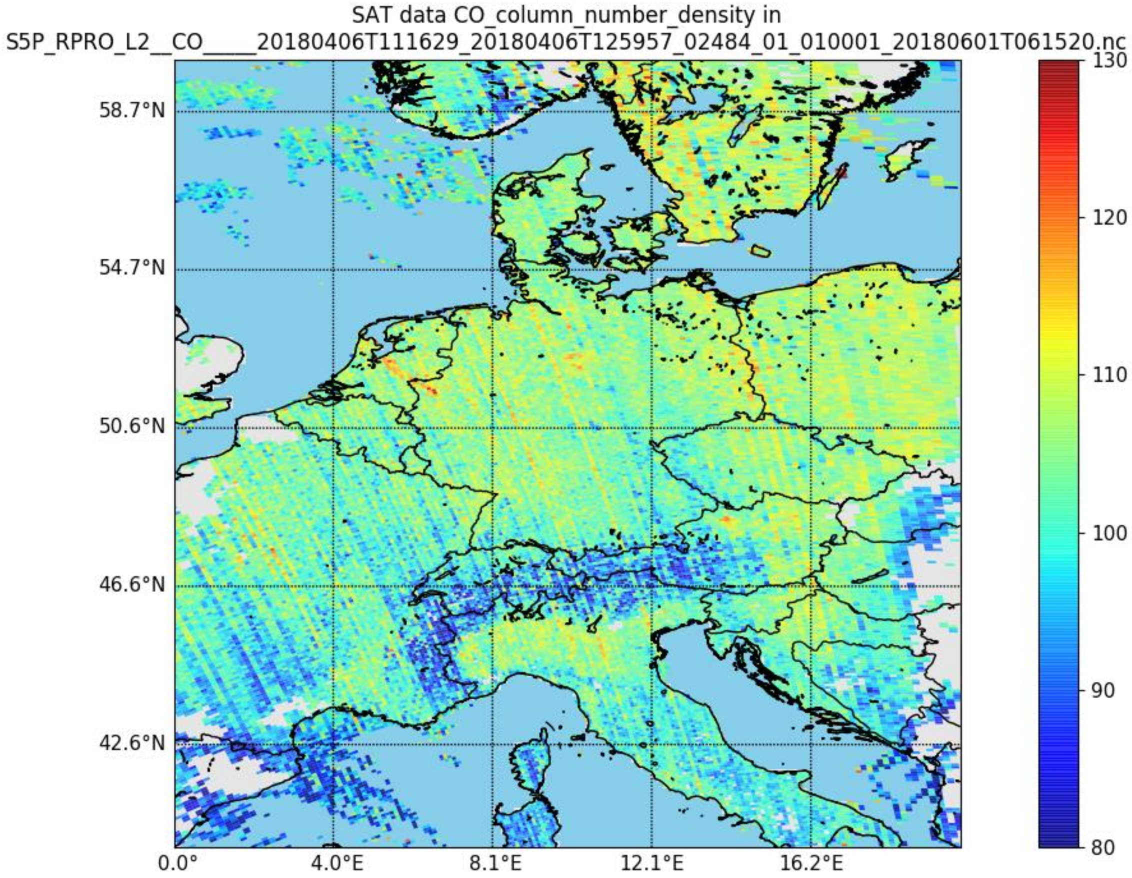


Figure 73: Example of stripe patterns in L2_CO OFFL CO column data along a S5P orbit above Europe.

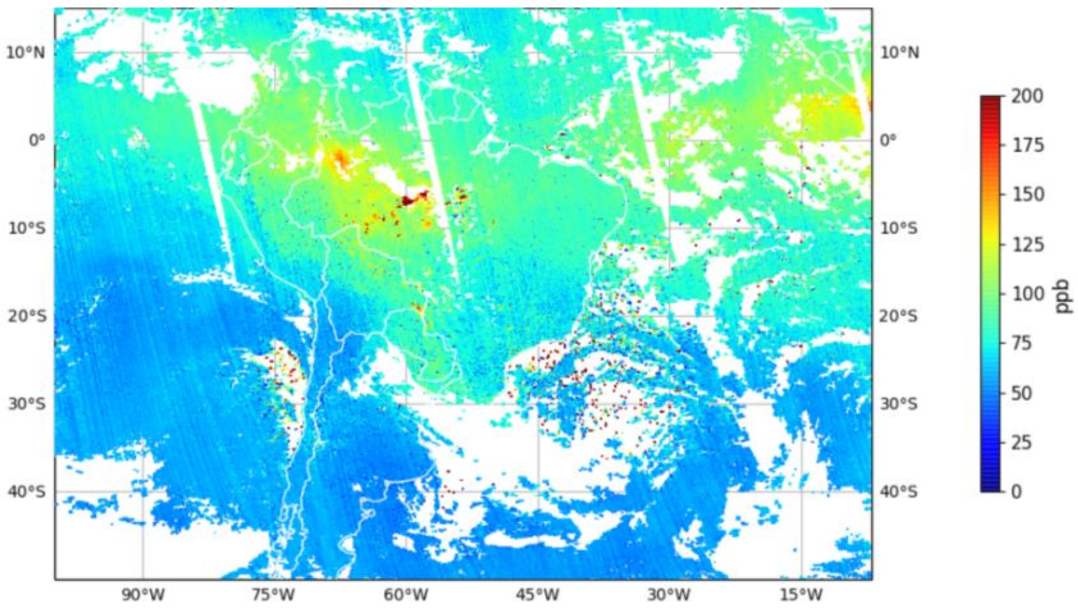


Figure 74: S5p OFFL xCO pixels measured on August 1, 2019, over South America and the Atlantic Ocean. Outlying pixels occurs (including negative values) in the South Atlantic Anomaly.

9.3.8 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.

9.4 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 75** confirms that the statistical quality indicators for both OFFL and NRTI since the processor change are very similar.

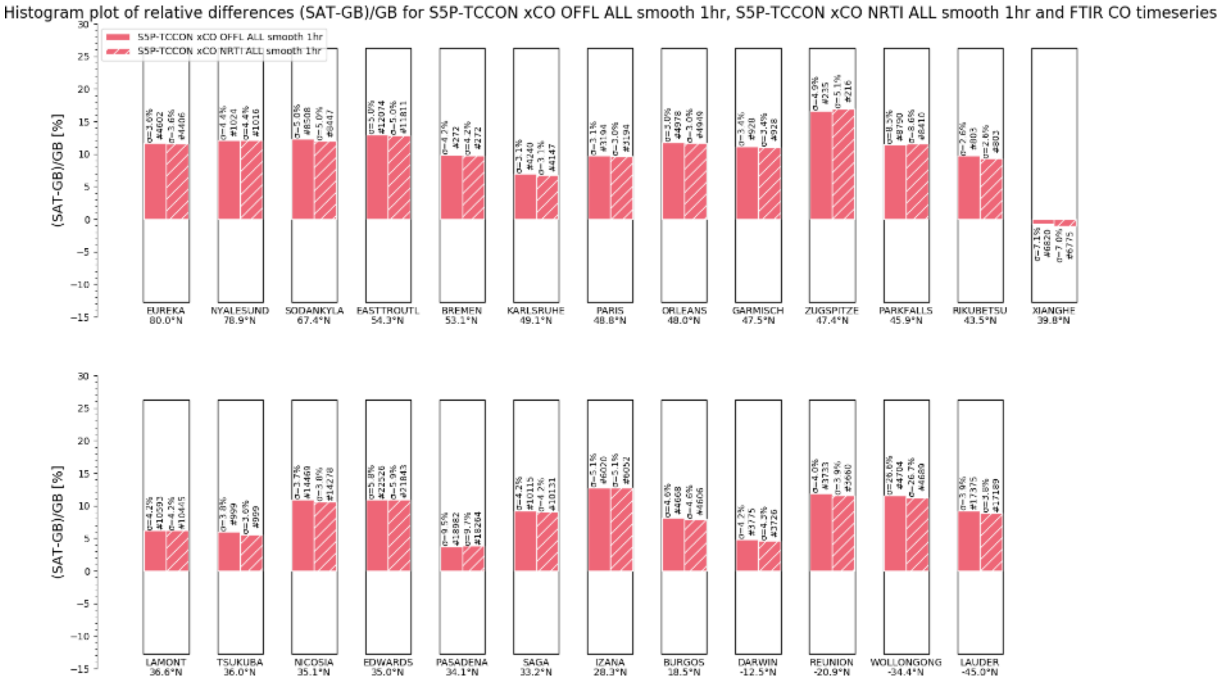


Figure 75: Comparison of relative biases against TCCON CO column data for the S5P L2_CO OFFL and NRTI data versions, from July 3 2019 through March 2021. The quality of both data sets is similar. Over this period the OFFL processor has produced data with a relative mean bias of 9.64% ± 5.59% and a correlation coefficient 0.93 with respect to TCCON data; the NRTI processor has produced data with a relative mean bias 9.50% ± 5.52% and a correlation 0.93.

10 Validation Results: L2_CH4

10.1 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**.

10.2 Validation approach

10.2.1 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) and the Total Carbon Column Observing Network (TCCON, https://tccondata.org). **Figure 67** displays the geographical distribution of the NDACC and TCCON sites. Near-infrared TCCON measurements provide calibrated methane column averaged (xCH_4) data with typical uncertainty values of 0.5% for the precision and 0.2% for the accuracy. Solar infrared NDACC measurements provide CH_4 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 measurements.

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_CH4 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 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, all available on 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 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.

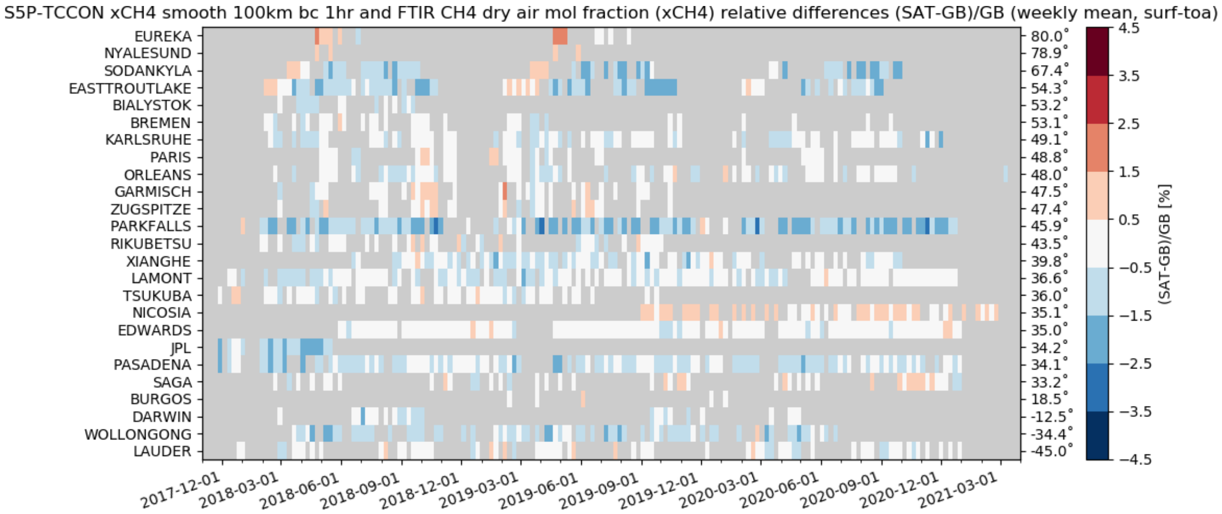
TROPOMI observations co-located with the TCCON measurements are found by selecting all filtered TROPOMI pixels within a radius of 100 km around each station and with a maximal time difference of 1h for TCCON and 3h for NDACC observations. The 1 hour interval can be justified by noting that TCCON instruments acquire only one type of spectra, while NDACC instruments are required to measure different types of spectra, making the number of CH_4 observations more sparse. In the comparison, the apriori in the TCCON and NDACC retrievals have been substituted with the S5P CH_4 apriori (Rodgers 2003). For NDACC data the method described in Rodgers (2003) is followed one step further and the FTIR CH_4 concentration profile (with the S5P prior substituted) is additionally smoothed with the S5P column averaging kernel. The validation procedure for both the NDACC and TCCON based comparisons includes an adaptation of the TROPOMI CH_4 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 December 2020. Routine validation is done using the Automated Validation Server of the MPC VDAF, the CH_4 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.

10.3.3 Bias

The systematic difference (the mean of all relative differences) between S5P L2_CH4 and TCCON dry air methane column data is on an average -0.68% (standard) and -0.25% (bias corrected), well within the mission requirements. Only at a few TCCON sites the bias is slightly higher than 1.5% for the standard S5P methane column.



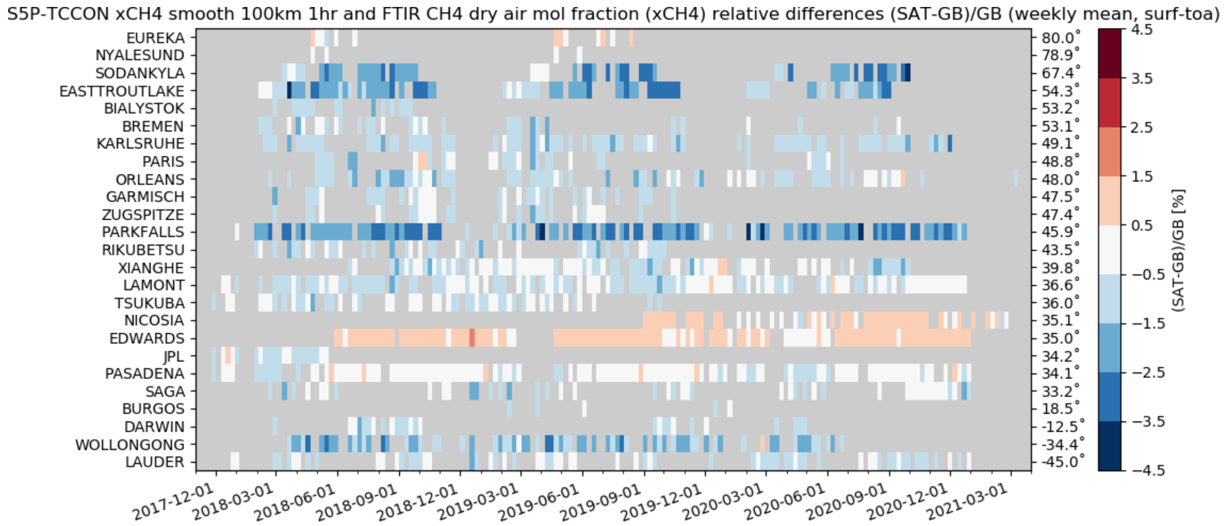


Figure 76: Mosaic plots of relative biases between S5P L2_CH4 RPRO+OFFL and ground-based CH_4 TCCON column data for the bias corrected (top) and standard (bottom) methane products. Over the November 2017 – March 2021 time period the plots do not show a clear meridian dependence or temporal change in the weekly averaged biases.

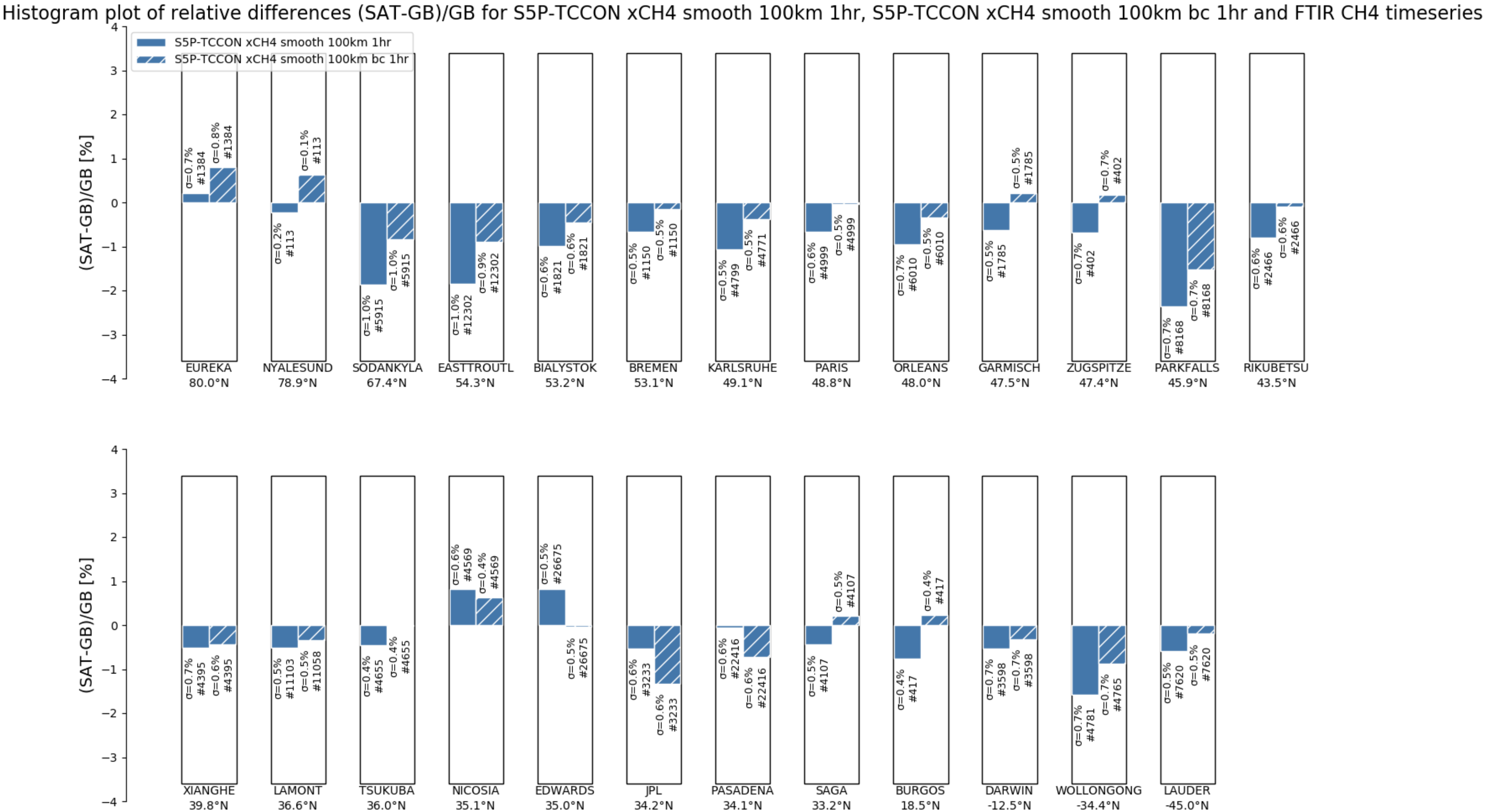


Figure 77: Chart of relative mean difference between S5P L2_CH4 and FTIR CH_4 column data at 25 TCCON sites within the time range November 2017 till March 2021. The sites are sorted with decreasing latitude. The relative mean difference of the standard xCH4 product slightly exceeds the mission requirements (bias below 1.5%) only at a few TCCON sites (i.e. Sodankylä, East Trout Lake, Park Falls and Wollongong). However, it never exceeds the mission requirements for the bias corrected product.

	S5P-NDACC xCH4 smooth 100km 3hr					S5P-NDACC xCH4 smooth 100km bc 3hr					
site	#	rel. NDACC std	correlation	rel diff bias(%)	rel diff std(%)	#	rel. NDACC std	correlation	rel diff bias(%)	rel diff std(%)	lat
EUREKA	43	1.9	0.44	0	1.22	43	1.5	0.46	0.82	1.23	80.1
NY.ALESUND	4	2.7	0.9	3.59	1.04	4	2.7	0.97	4.56	1	78.9
THULE	67	1	0.74	3.41	0.8	67	0.9	0.76	4.46	0.78	76.5
KIRUNA	121	1	0.14	-1.2	1.42	121	1	0.14	-0.21	1.46	67.8
SODANKYLA	124	1.1	0.37	-0.99	1.14	124	1.1	0.41	0.06	1.12	67.4
HARESTUA	48	1.5	0.41	-0.33	1.06	48	1.5	0.42	0.81	1.05	60.2
ST.PETERSBURG	104	1.2	0.4	-0.74	0.95	104	1.2	0.51	0.22	0.86	59.9
BREMEN	51	1.9	0.55	0.74	1.41	51	2	0.56	1.28	1.4	53.1
KARLSRUHE	173	0.8	0.78	-0.35	0.52	173	0.9	0.8	0.33	0.47	49.1
GARMISCH	17	1.1	0.73	-0.39	0.54	17	1.1	0.76	0.48	0.52	47.5
ZUGSPITZE	60	1	0.67	-0.03	0.63	60	1	0.69	0.8	0.62	47.4
JUNGFRAUJOCH	61	1.1	0.59	-0.86	0.84	61	1	0.62	-0.19	0.83	46.6
TORONTO	122	2.7	0	0.93	2.95	122	2.6	-0.04	1.52	3.02	43.6
RIKUBETSU	36	2.1	0.23	0.82	2.08	36	1.9	0.1	1.49	2.24	43.5
BOULDER.CO	121	1.3	0.62	1.87	0.83	121	1.2	0.51	1.48	0.94	40
ALTZOMONI	58	1.1	0.16	2.53	0.82	58	1	0.36	2.36	0.74	19.1
WOLLONGONG	162	0.9	0.67	-0.79	0.79	162	0.9	0.7	-0.07	0.75	-34.4
LAUDER	94	1.3	0.7	-0.2	0.9	94	1.4	0.66	0.2	0.96	-45
ARRIVAL.HEIGHTS	38	0.8	0.64	1.56	0.83	38	0.8	0.66	2.56	0.82	-77.8
		1.4	0.51	0.5	1.09		1.4	0.53	1.21	1.1	

Table 12 – Overview of statistical quality indicators for the co-located S5P and NDACC time series. Due to the lower accuracy of the NDACC data, conclusions can be drawn on precision only (std on the rel. diff.).

10.3.4 Dispersion

The 1σ spread of the relative difference (between the S5P and the TCCON 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 sites are indicated in **Figure 77**.

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) is hard to evaluate: at high latitude sites 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).

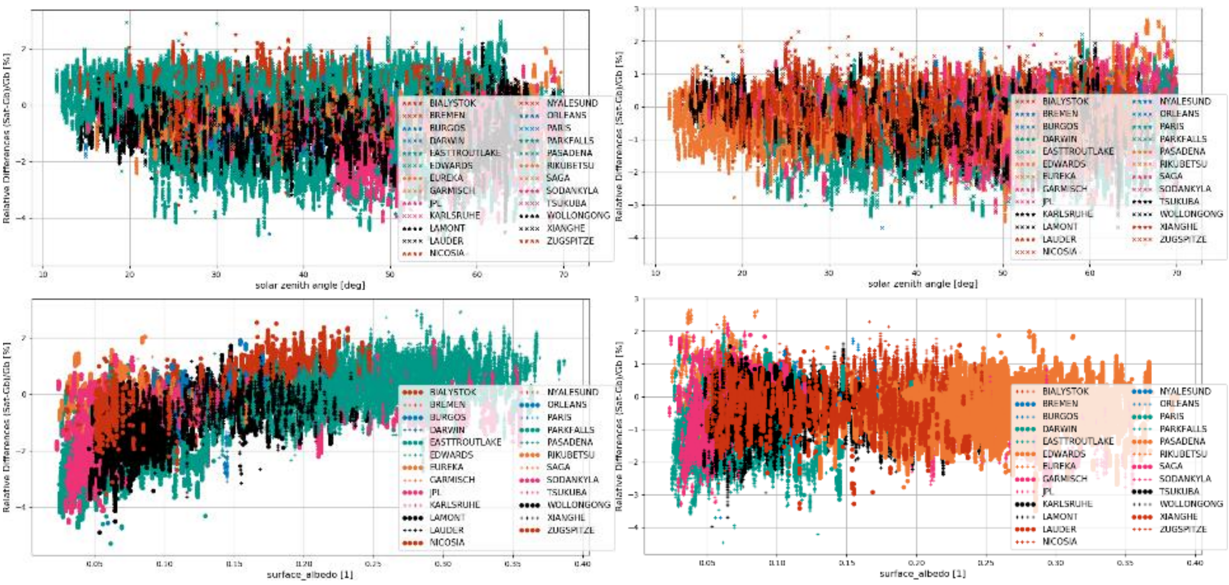


Figure 78: Dependence of the S5P-TCCON relative difference on solar zenith angle (top) and surface albedo (bottom). The left column shows the standard S5P product and the right column the bias corrected S5P product. The bias correction removes the surface albedo dependence of the standard S5P product.

10.3.6 Short term variability

For all the NDACC and TCCON stations, short scale temporal variations in the CH_4 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 79.** At some sites the correlations are very low (e.g. Sodankylä, Darwin). This is probably due to the `qa_value` filtering which, at some sites, does not filter all bad pixels, see also Section 10.3.8.

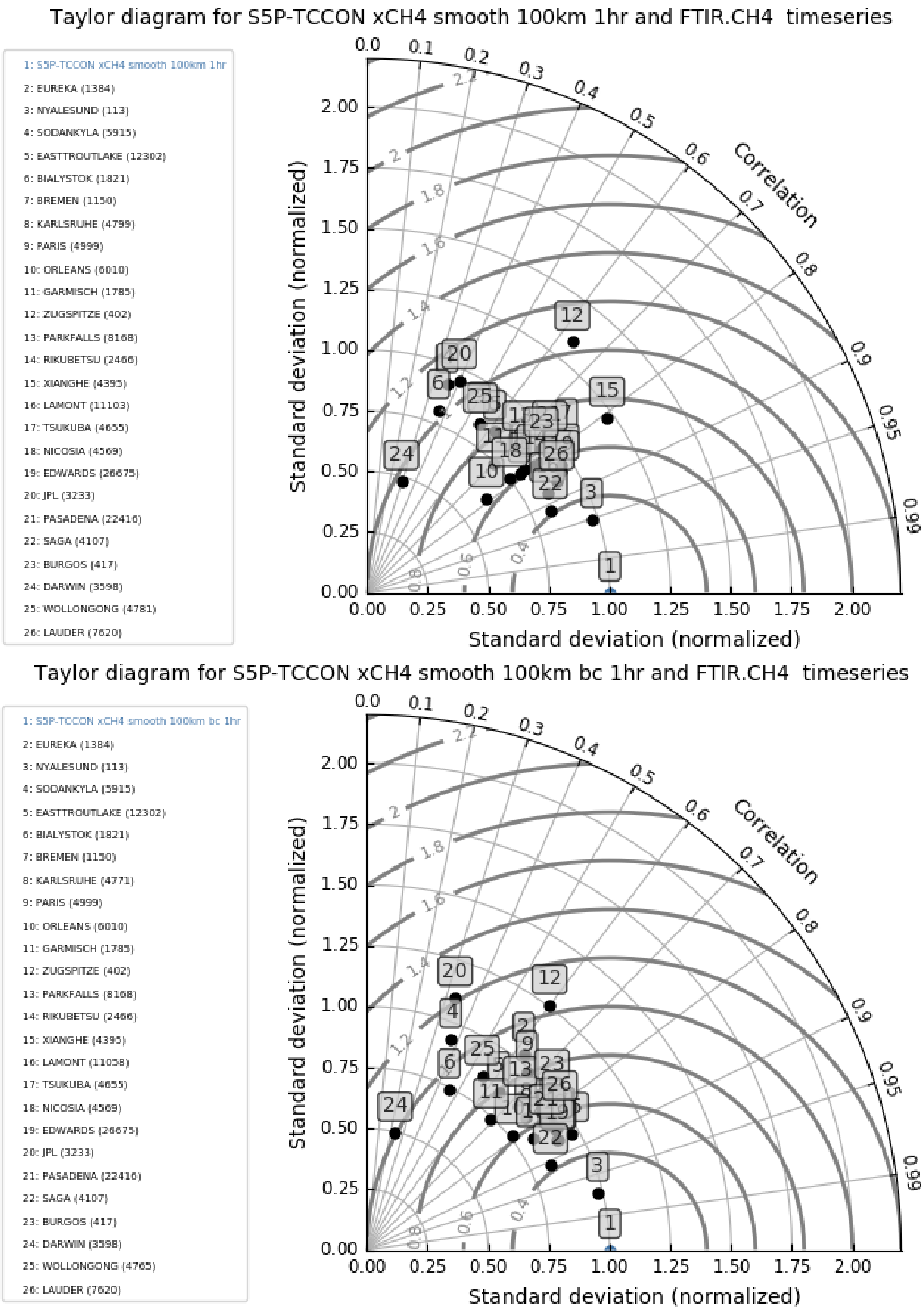


Figure 79: Taylor diagrams for differences between S5P L2_CH4 RPRO/OFFL and TCCON methane column data: standard (top) and bias corrected (bottom) S5P methane columns. At almost all sites the variability of the S5P column data is higher compared to the variability in the TCCON data.

10.3.7 Geographical patterns

Single TROPOMI overpasses show stripes of erroneous CH_4 values in the flight direction (see **Figure 80** left). For orbits before orbit 7644 (April 5 2019) not all pixels above inland water are filtered with the `qa_value` flag, see **Figure 80** (right, above Caspian Sea).

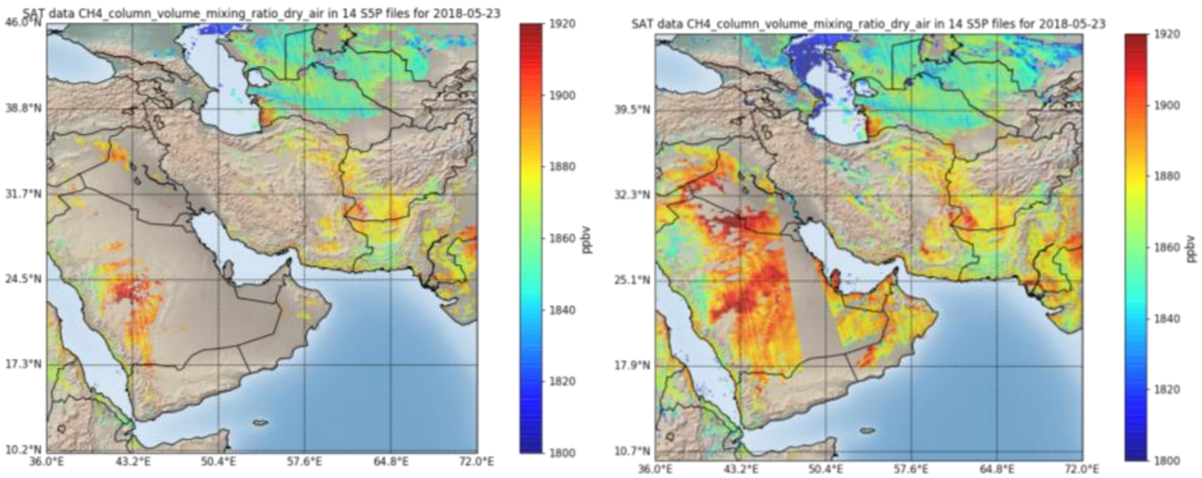


Figure 80: Maps showing XCH_4 concentrations above the Middle East measured by S5P on May 23 2018. The left panel shows all available ground pixels, the right panel shows only pixels with qa_value > 0.5. The left panel shows the presence of stripes in the flight direction and the right panel shows the presence of filtered pixels above inland water (Caspian Sea).

10.3.8 Change in Cloud Mask

Since March 11 2020 the cloud data used for the CH_4 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 XCH_4 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 81**). 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 01.04.00 implemented since November 29, 2020.

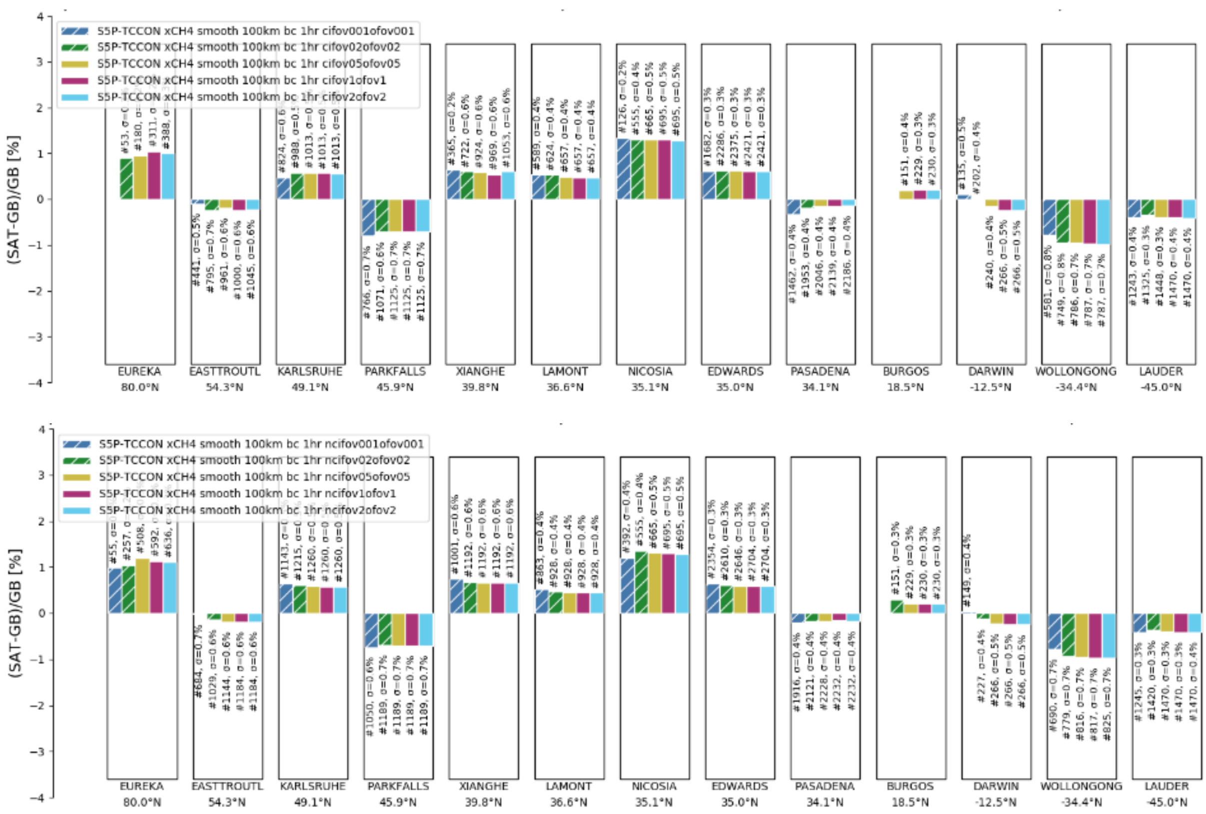


Figure 81: charts of relative mean difference between S5P L2_CH4 and FTIR CH_4 column data at 13 TCCON sites within the time range 11 March 2020 till 19 July 2020. The sites are sorted with decreasing latitude. Top panel plot is for the old CF definition and the bottom panel plot is for the new CF definition. Different colours represent the different CF threshold applied for the filtering.

10.3.9 Other features

Filtering on `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 82**).

Outlying methane values are observed along coastline regions or mountain regions, see for example Greenland in **Figure 83**.

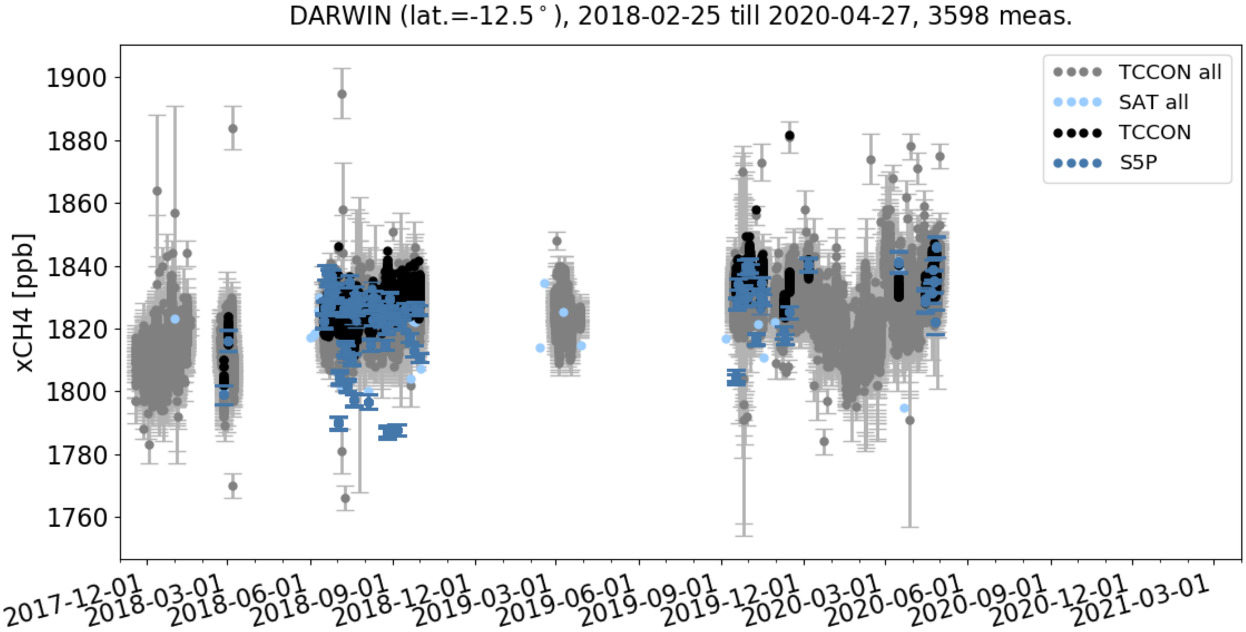


Figure 82: S5P L2_CH4 XCH$_4$ time series over Darwin where low values of XCH$_4$ are observed for several days.

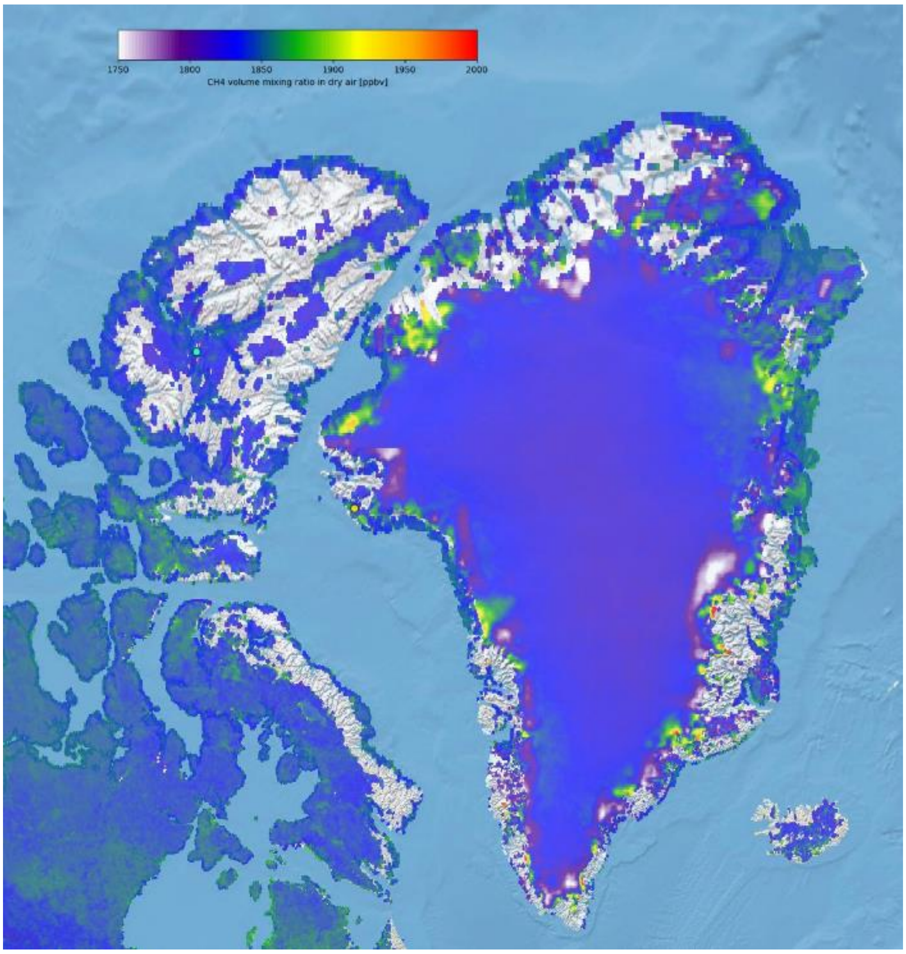


Figure 83: Map showing XCH4 concentrations above Greenland and parts of North America for three years averaged data with pixels with qa_value > 0.5. The XCH4 pixels in Greenland shows outliers along the mountain and coastline regions. Similar features are observed at the Antarctic.

11 Validation Results: L2_CLOUD

11.1 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 11.3 concentrates on the validation of the L2_CLOUD OFFL product while Subsection 11.4 demonstrates evidence that NRTI and OFFL data do not differ significantly and that their respective validations yield similar conclusions before the processor switch.

11.2 Validation approach

11.2.1 Ground-based networks

CLOUDNET lidar/radar data

S5P TROPOMI L2_CLOUD cloud data have been routinely compared at 21 ground-based stations (**Table 13**) 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 site, and the cloud is not multilayered 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 site, the site 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.

Station	Location	Network	Latitude (N)	Longitude (E)
Ny-Ålesund	Svalbard	CLOUDNET	78.932	11.921
Summit	Greenland	NOAA/ARM	72.60	-38.42
Andoya	Norway	ARM	69.14	15.68
Kenttarova	Finland	CLOUDNET	67.987	24.23
Hyytiala	Finland	CLOUDNET	61.84	24.29
Norunda	Sweden	CLOUDNET	60.09	17.48
Mace Head	Ireland	CLOUDNET	53.325	-9.9
Lindenberg	Germany	CLOUDNET	52.211	14.13
Leipzig	Germany	CLOUDNET	51.35	12.43
Chilbolton	United Kingdom	CLOUDNET	51.145	-1.437
Jülich	Germany	CLOUDNET	50.909	6.413
Palaiseau	France	CLOUDNET	48.713	2.208
Munich	Germany	CLOUDNET	48.15	11.57
Schneefernerhaus	Germany	CLOUDNET	47.42	10.98
Bucharest	Romania	CLOUDNET	44.35	26.03
Potenza	Italy	CLOUDNET	40.6	15.72
Graciosa	Azores	ARM	39.092	-28.026
Granada	Spain	CLOUDNET	37.16	-3.605
Iquique	Chile	CLOUDNET	-20.54	-70.18
Maido	La Reunion	CLOUDNET	-21.08	55.38
Villa Yacanto	Argentina	ARM	-32.13	-64.73

Table 13 – List of ground-based stations providing the cloud classification data product, and used in this study: 17 CLOUDNET sites and 4 ARM (Atmospheric Radiation Measurement) sites. Data is collected through EVDC.

11.2.2 Satellites

MODIS, NPP VIIRS and OMI O_2-O_2

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 product[1] 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-atmos.gsfc.nasa.gov/sites/default/files/ModAtmo/EOSSNPPCloudOpticalPropertyContinuityProductUserGuidev1.pdf . Those operational publicly available data might have some differences compared to the limited data provided by the NASA group directly to DLR in 2018.

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).

11.2.3 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 sites. 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 O_2 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

S5P CLOUD pixels and S5P FRESCO pixels covering CLOUDNET sites were extracted. For both CLOUD and FRESCO, only pixels with qa_value > 0.5, and with CF_rescaled > 0.1 were kept, and common overpasses were considered.

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 radiomatric cloud fraction) here.

11.2.4 Field campaigns and modelling support

None for this report.

11.3 Validation of L2_CLOUD OFFL

11.3.1 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>75º), 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.

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. 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-vdaf.tropomi.eu.

The validation vs. CLOUDNET ground-based data uses S5P L2_CLOUD RPRO+OFFL 01.01.07-02.01.04 data. This covers the time period from 2018-04-30 till May 2021. CLOUDNET data from 21 sites were considered in this analysis. Note that the station Kenttarova did not provide valid co-locations in the CLOUDNET measurement period (20191027-20191225).

11.3.3 Radiometric cloud fraction (L2_CLOUD CAL & L2_CLOUD CRB)

11.3.3.1 Bias

Comparison with other satellites and alternative algorithm S5P FRESCO

Figure 84 presents the latitudinal variation of the zonal mean of the radiometric cloud fraction (left), scaled radiometric cloud fraction (middle) and cloud (top) height (right) of the cloud products S5P OFFL CLOUD (CAL and CRB), S5P OFFL FRESCO and Aura/OMI OMCLDO2 at two different days: 2020/07/09 and 2020/07/18. I.e., this is before and after the major version upgrade, at 2020/07/13, orbit 14239, there was a major version upgrade from 1.1.8 to 2.1.3. Note that there was no spatial pixel co-location performed, which can affect results.

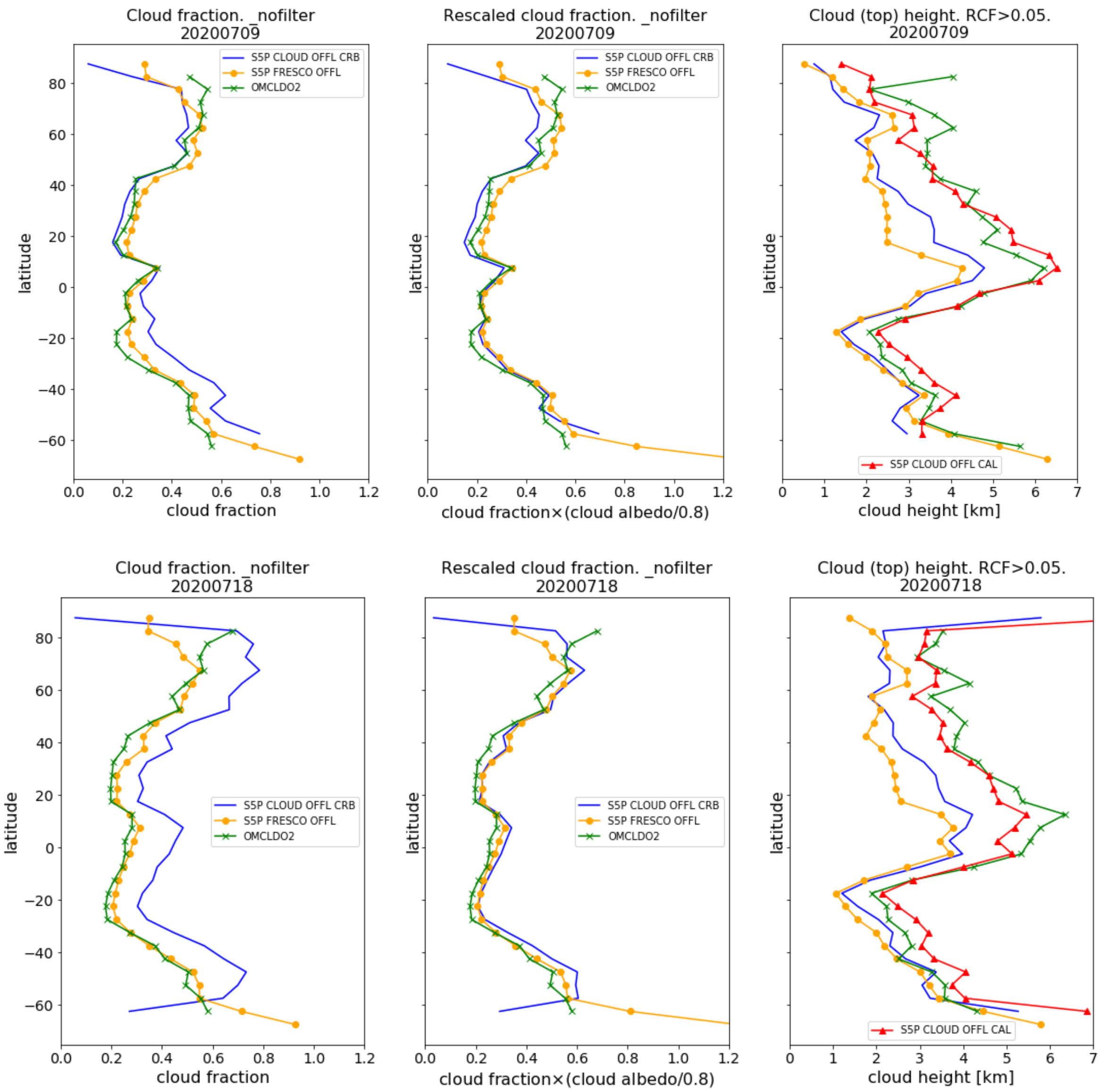


Figure 84: Latitudinal variation of zonal means of cloud fraction (left), scaled cloud fraction (middle) and cloud (top) height (right) of the cloud products S5P OFFL CLOUD (CAL and CRB), S5P OFFL FRESCO 1.3 and Aura/OMI OMCLDO2. Top: at 2020/07/09 (UPAS version = 1.1.8). Bottom: at 2020/07/18 (UPAS version = 2.1.3).

Several other days in July were investigated (up to 2020/07/19) with similar conclusions:

- The CRB radiometric cloud fraction is enhanced considerably after the UPAS upgrade, becoming much higher than those of S5P FRESCO and OMCLDO2. Note however that these cloud fractions cannot be directly compared as they have different assumptions on cloud albedo.
- Scaled radiometric cloud fraction (i.e., (cloud fraction x cloud albedo)/0.8, shorthand sRCF) are better comparable quantities. The scaled radiometric cloud fraction of CLOUD CRB is close to that of S5P FRESCO and OMCLDO2..Note that the fact that scaled radiometric cloud fractions still match, implies that CRB cloud albedo is decreased with the UPAS upgrade. At the more extreme latitudes S5P CLOUD CRB and S5P FRESCO diverge, likely due to a different treatment of snow-ice cover.

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 S5P CLOUD products.

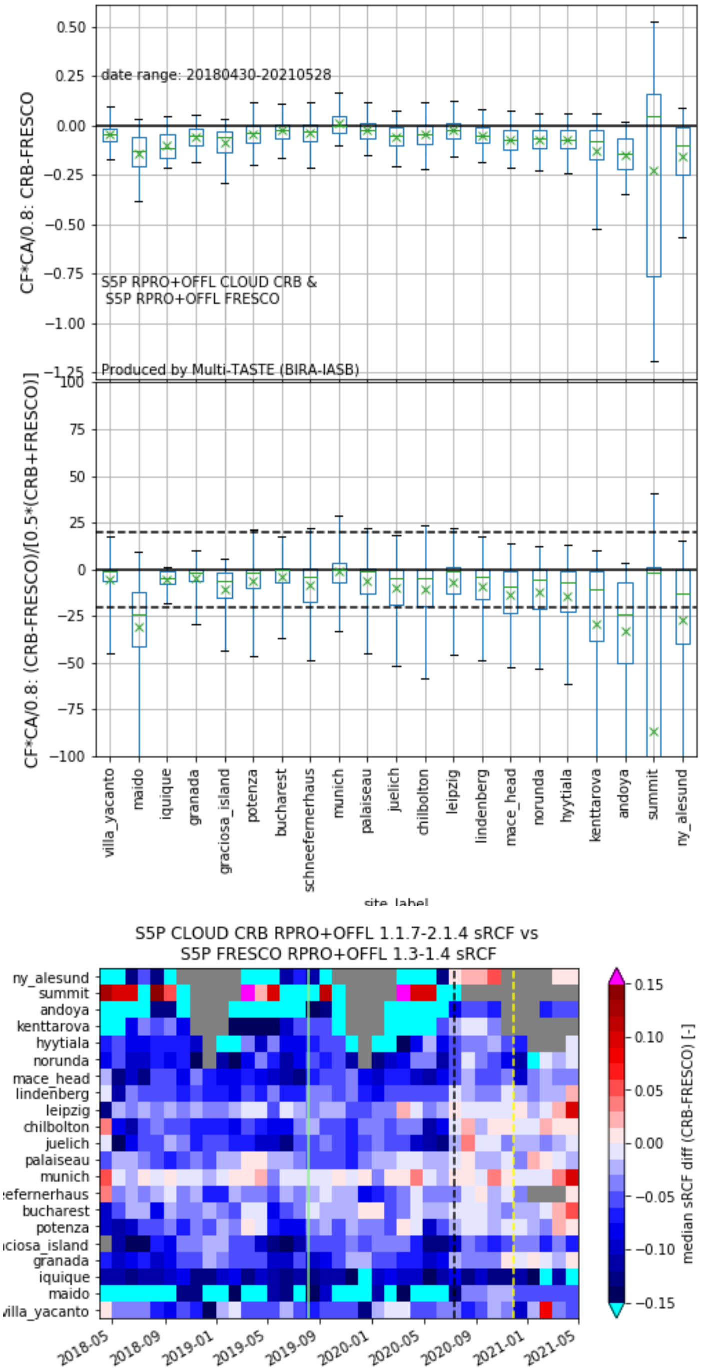
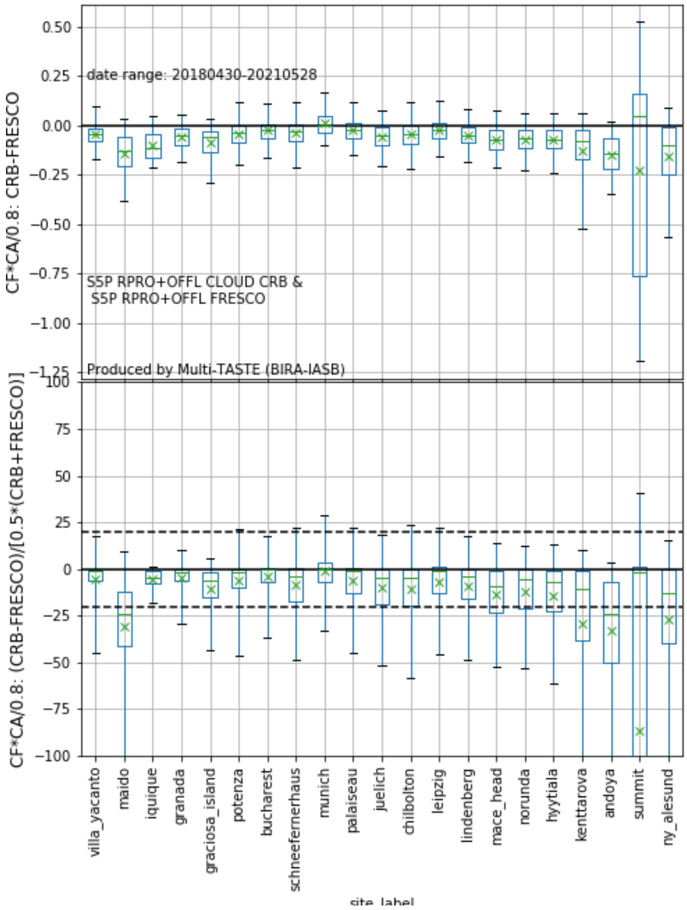


Figure 85Figure 85 presents boxplots of the difference (top) and normed relative difference (center) between rescaled cloud fractions of CLOUD CRB and FRESCO at the CLOUDNET sites, as well as the evolution of the monthly median difference over the CLOUDNET sites. Overall, mean and median difference are negative (CRB lower than FRESCO) but in most cases within the 20% bias requirement.

The strong deviations at the Northern sites 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.

There is an important impact of version upgrades. With the upgrade from CLOUD v1 to v2, the median difference becomes lower for most sites. In particular, the strongly negative difference at the Northern sites Ny-Alesund becomes positive. This can at least partly be attributed to the better surface albedo treatment of this complicated sea/land/snow/ice site, and to a more strict snow/ice filtering. Strong reductions at Andoya and Maido are also visible.



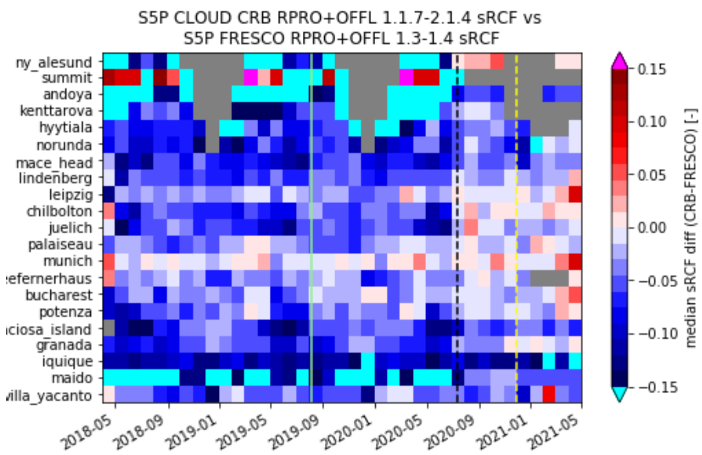


Figure 85. Top. Box plots of S5P CLOUD CRB OFFL cloud fraction minus S5P FRESCO OFFL cloud fraction, after both have been rescaled to a cloud albedo of 0.8. Middle: same but for the normed relative difference. Bottom: Monthly median of scaled cloud fraction difference, per month and per site. Indicated are the pixel size switch (2019/08/06), the CLOUD v1 to v2 upgrade (2020/07/13) and the FRESCO v1.3 to 1.4 upgrade (2020/11/29).

11.3.3.2 Dispersion

Figure 86 (left) presents the evolution of the difference dispersion (½ IP68 (CRB minus FRESCO)) over the CLOUDNET sites, per month. **Figure 86** (right) presents the overall dispersion per site. . At most sites, the ½ IP68 slightly exceeds the dispersion requirement of 0.05, while there is a high dispersion at the high-latitude sites Kenttarova, Andoya, Summit and Ny-Alesund (

Figure 85, **Figure 86**). This is likely due to a different treatment of snow/ice by both algorithms (see also section 11.3.3.3). Note also the high dispersion at Maido (Reunion Island). The dispersion at Ny-Alesund is reduced with the upgrade from CLOUD v1 to CLOUD v2. Note that after this upgrade, there is no more data from Summit, due to a more strict snow/ice filtering.

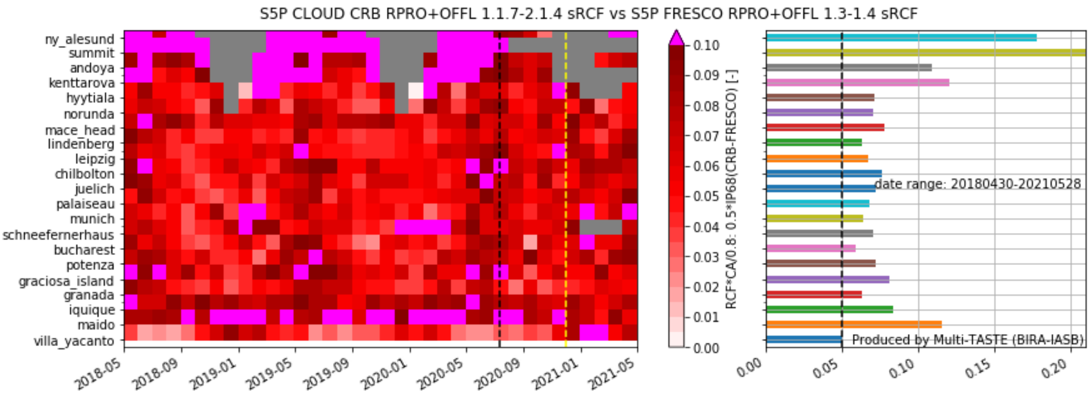


Figure 86. Left. Monthly ½ IP68 of scaled cloud fraction difference of CLOUD OFFL vs FRESCO OFFL, per month and per site. Indicated are the pixel size switch (2019/08/06), the CLOUD v1 to v2 upgrade (2020/07/13) and the FRESCO v1.3 to 1.4 upgrade (2020/11/29). Right: ½ IP68 of scaled cloud fraction difference (CLOUD CRB-FRESCO), per site. Note that the dispersion at the site Summit is off-scale and extends to ~0.6 km.

11.3.3.3 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.

11.3.3.4 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 is believed to be minimized with the upcoming L1b data version 2 update which contains a degradation correction.

11.3.3.5 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.

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

11.3.4.1 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. S5P CLOUD CRB CH was compared with S5P FRESCO CH, over the CLOUDNET sites (**Figure 87**). At most sites, a small bias is encountered between 0 and -0.2 km (mean and median difference; CRB lower than FRESCO) or ~-10% (median normed relative difference). This is within the 20% bias requirement. Clear exceptions (CRB higher than FRESCO) occur at the sites Iquique, Ny-Ålesund and Summit. This boxplot however hides the important impact of version upgrades (**Figure 87**, bottom). 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-Alesund becomes negative. The issue of exceptionally high CRB values is resolved with the upgrade to CLOUD v2. 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.

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 88 to **Figure 90** 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 1.3 and FRESCO 1.4, separately for low and high clouds (separation at CLOUDNET CTH=4km).

Conclusions are as follows:

- CLOUD CAL CTH is slightly above the CLOUDNET cloud mean height (CMH) for high clouds while for low clouds it is mostly in between CLOUDNET CTH and CMH,
- CLOUD CRB CH is in between CLOUDNET CMH and CLOUDNET CBH for high clouds, while for low clouds it is close to the CLOUDNET CBH.

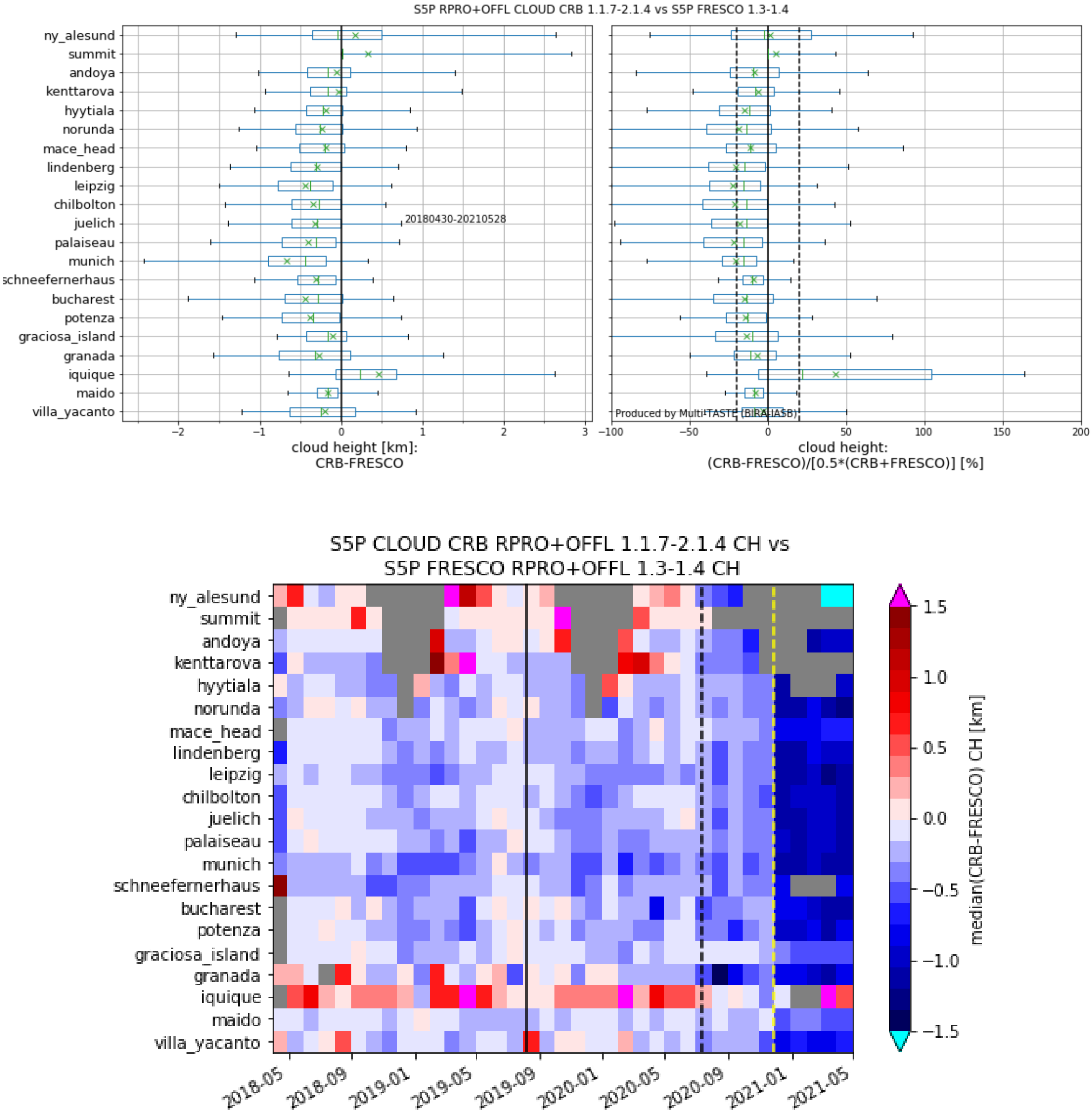


Figure 87: Top. Boxplots of S5P CLOUD CRB CH minus S5P FRESCO CH (left) and of the normed difference (right). Bottom. Monthly median difference per station.

- The exceptional overshoot of CLOUD cloud height value at Ny-Alesund disappears with the upgrade from CLOUD v1 to CLOUD v2.
- For high clouds, FRESCO CH is slightly below CLOUDNET CMH before its upgrade to 1.4, and slightly above after this upgrade. 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 sites Norunda and Hyytiala, 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-Alesund) sites present a seasonal cycle in the bias, with high FRESCO CH early 2020 and 2021 (see Figure 91). 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.

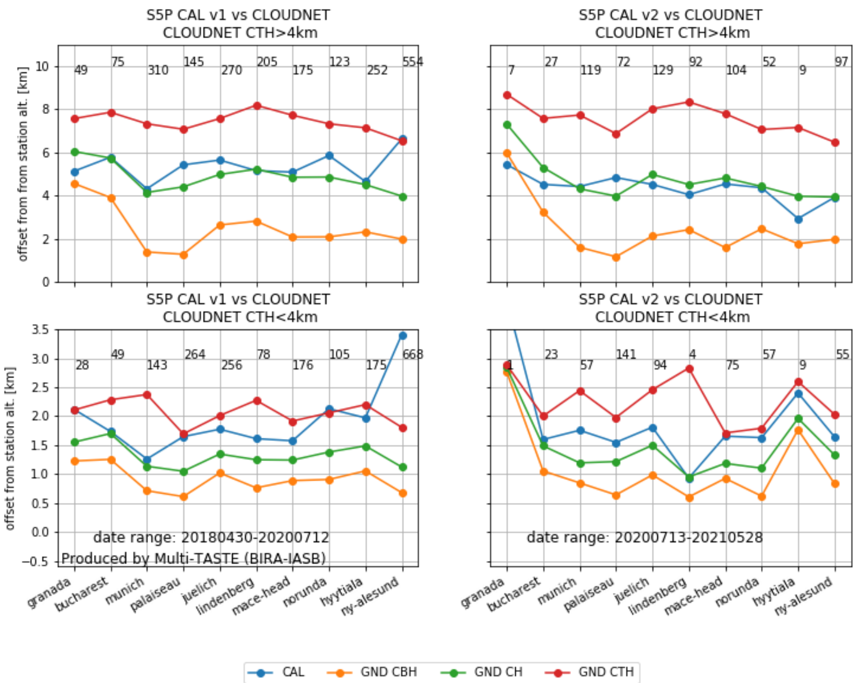


Figure 88. Mean values of the offset from the CLOUDNET sensor altitude, per site of (i) S5P CLOUD OFFL CAL cloud top height, (ii) CLOUDNET cloud base height, (iii) CLOUDNET cloud mean height (calculated from cloudy altitude bins), (iv) CLOUDNET cloud top height. Top/bottom row panels: high and low clouds respectively (split at CLOUDNET CTH=4km). Left/right column panels: CLOUD v1 and CLOUD v2 respectively. Only sites with data for both versions are included. Data count per site is given on the plot.

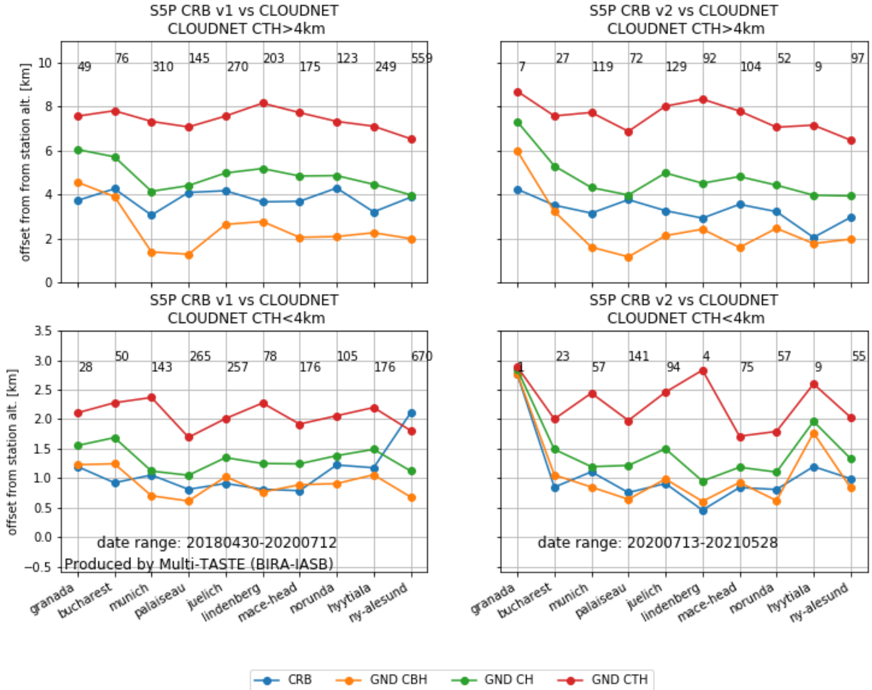


Figure 89. Same as **Figure 88** but for S5P CLOUD OFFL CRB CH instead of S5P CLOUD OFFL CAL CTH.

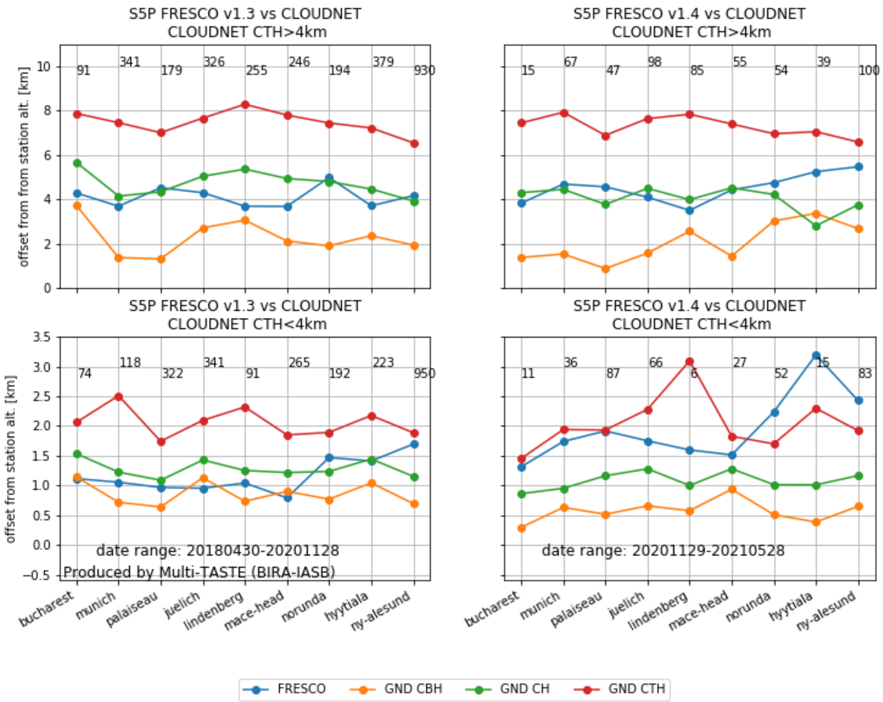


Figure 90. Same as **Figure 88** but for S5P OFFL FRESCO CH instead of S5P CLOUD OFFL CAL CTH, and the version split at FRESCO v1.3 and FRESCO v1.4.

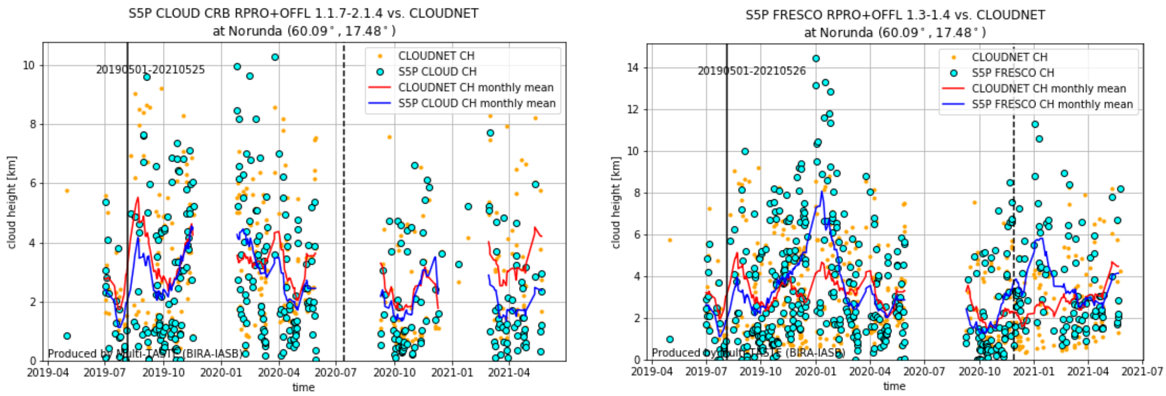


Figure 91. Left. Time series of S5P CLOUD CRB OFFL CH and CLOUDNET CMH. Right. Time series of S5P FRESCO OFFL CH and CLOUDNET CMH. Note the peak in FRESCO CH in early 2020 and 2021, while CLOUD CRB has almost no data in this period.

L2_CLOUD CAL cloud top height is generally below the CLOUDNET cloud top height. This can be seen in **Figure 92**, which presents boxplots per site of absolute-scale and relative difference (top panels) and of the monthly median difference (bottom panel). The initially positive bias at Ny-Alesund becomes negative with the upgrade to CLOUD v2. High negative biases at Granada and Lindenberg (exceeding 3 km) and a more negative bias at Juelich 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., 2020) 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-site 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 14**). Note that for v2, the number of sites is limited and collocations at some sites are sparse.

Table 14. Statistics, calculated over the sites of S5P CLOUD OFFL CAL CTH-CLOUDNET CTH, S5P CLOUD OFFL CRB CH-CLOUDNET CH and S5P FRESCO OFFL CH-CLOUDNET CH. Measurements up to 2021/05/28 are included.

CTH: CAL -CLOUDNET	**#sites**	**Med(meddiff)**	**Med(medreldiff)**	**Med(1/2IP68)**
CLOUD v1; high	20	-2 km	-28%	2.1 km
CLOUD v1; low	20	-0.3 km	-16%	0.8 km
CLOUD v2; high	10	-3 km	-42%	1.9 km
CLOUD v2; low	10	-0.4 km	-17%	0.5 km
CH: CRB-CLOUDNET				
CLOUD v1; high	20	-0.9 km	-21%	1.6 km
CLOUD v1; low	20	-0.5 km	-39%	0.5 km
CLOUD v2; high	10	-1.0 km	-27%	1.5 km
CLOUD v2; low	10	-0.5 km	-38%	0.4 km
CH: FRESCO-CLOUDNET				
FRESCO v1.3; high	21	-0.8 km	-15%	2 km
FRESCO v1.3; low	21	-0.4 km	-21%	0.6 km
FRESCO v1.4; high	9	0.4 km	9%	1.7 km
FRESCO v1.4; low	9	0.6 km	55%	0.5 km

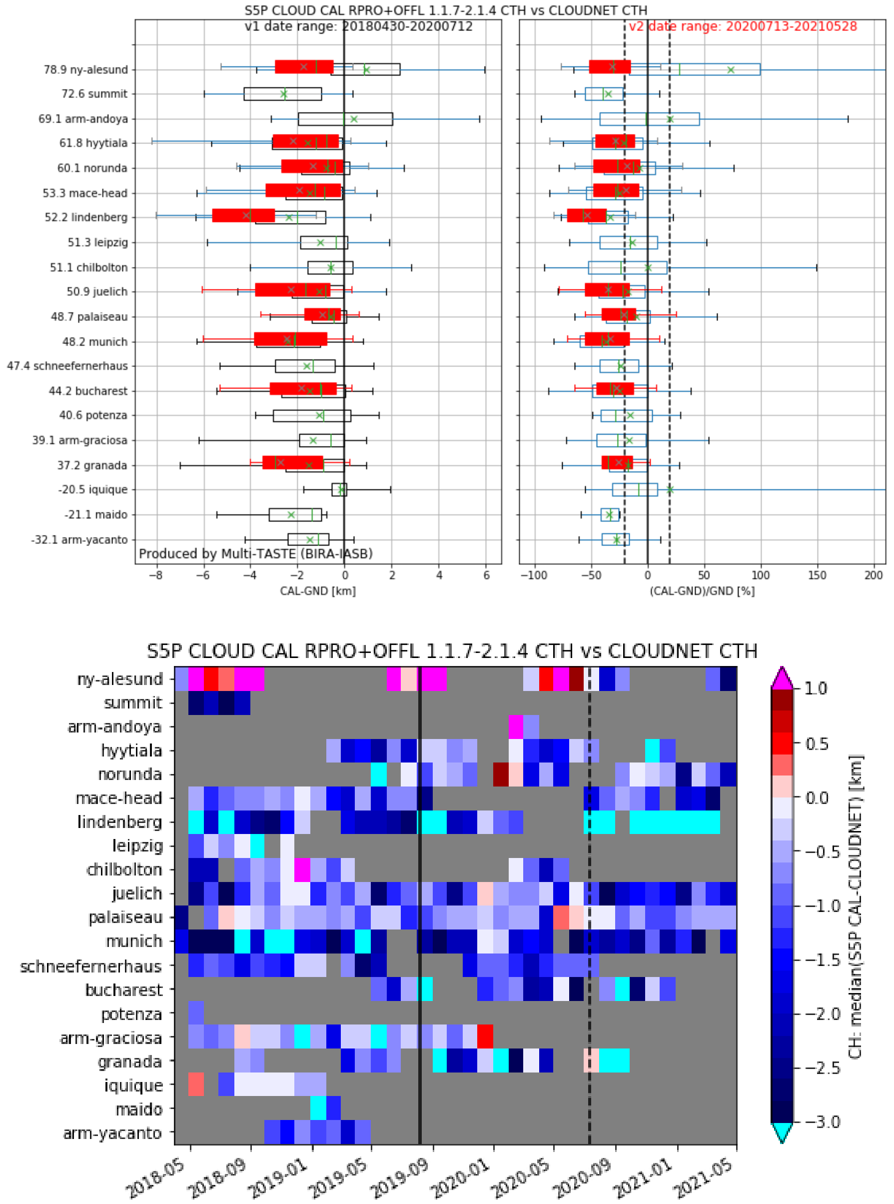


Figure 92: Top panel: Boxplots of S5P L2_CLOUD CAL CTH minus CLOUDNET CTH, per site. Boxplot conventions: box bounds are at first and third quartile. Green line is median. Whiskers are at 5 and 95 percentiles. Green cross is mean. Right panel: Similar as upper panel but now for the relative difference. Unfilled boxplots are for CLOUD v1, filled boxplot for CLOUD v2. Sensing date range is indicated on the plot. Bottom panel: monthly median of CLOUD CAL CTH-CLOUDNET CTH, per site. Pixel size switch (2019/08/06) and UPAS version change (2020/07/13) are indicated on the plot.

L2_CLOUD CRB cloud height is generally below the CLOUDNET cloud mean height. This can be seen in Figure 93, which presents boxplots per site of absolute-scale and relative difference (top panels) and of the monthly median difference (bottom panel). The initially positive bias at Ny-Alesund 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 93). While there are no CLOUDNET co-locations available for CLOUD CRB v2, we note that the improvement 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-site 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% (rounded to 5%) are obtained for high clouds, and -0.5 km and -40% for low clouds (see **Table 14**). Note that the number of sites used for version 2 is limited.

Figure 94 displays boxplots of the difference and relative difference between S5P FRESCO CH and CLOUDNET CH, and the monthly medians per site. 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 -0.8 km (-15%) for high clouds, and -0.4 km (-20%) for low clouds. For version 1.4, bias (relative bias) are now positive: 0.8 km (15%) for high clouds, and 0.6 km (55%) for low clouds.

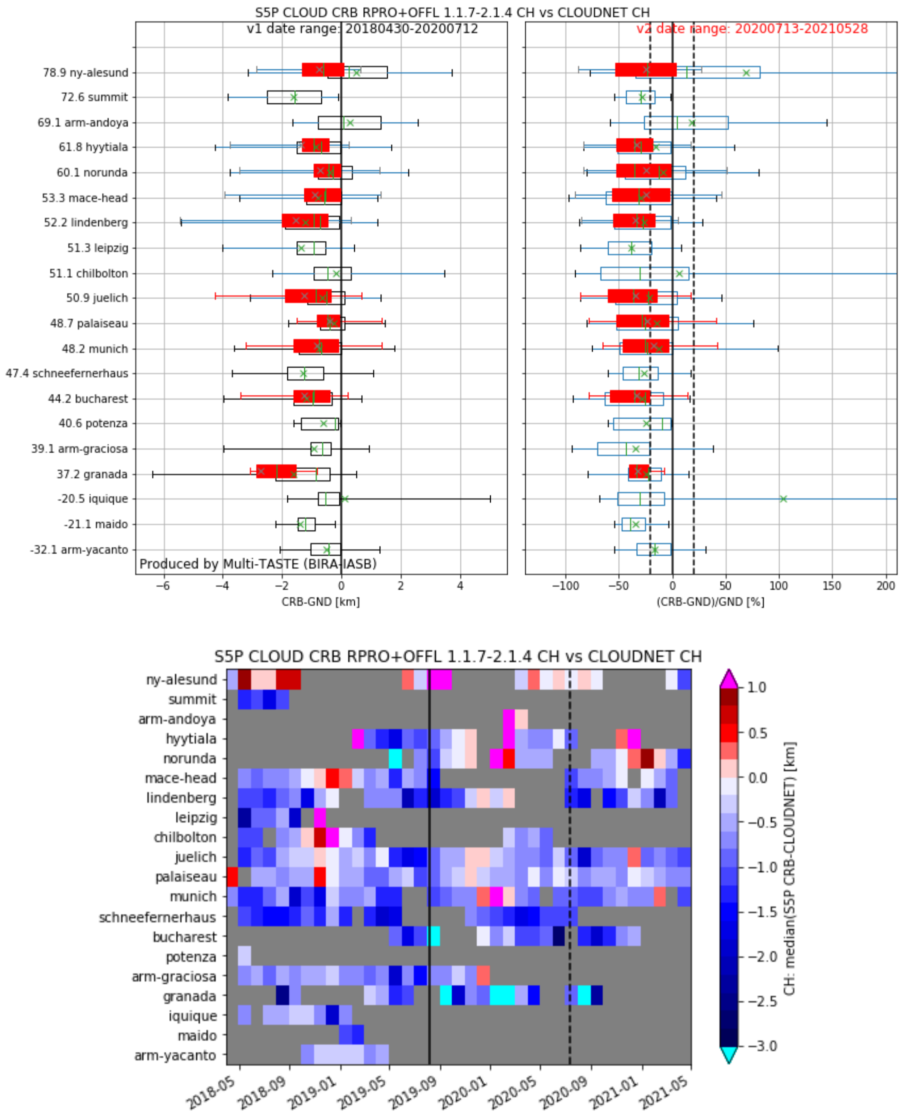


Figure 93: Upper panel: Boxplots of S5P L2_CLOUD CRB RPRO+OFFL CH minus CLOUDNET CH (upper left) and of the relative difference (upper right), per site. The same conventions as for **Table 14** apply. Sensing time range is indicated on the figure. Bottom panel: monthly median of CLOUD CRB CH-CLOUDNET CH, per site. Pixel size switch (2019/08/06) and UPAS version change (2020/07/13) are indicated on the plot.

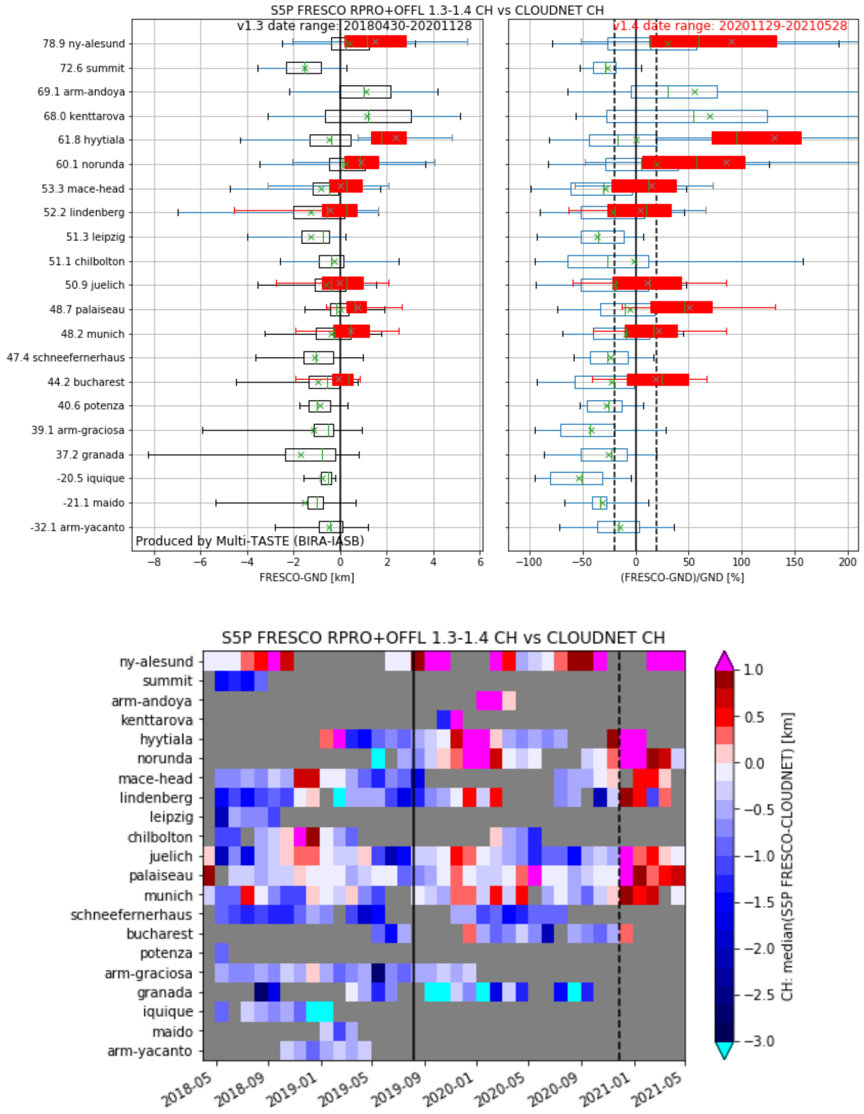


Figure 94: Upper panel: Boxplots of S5P L2_CLOUD FRESCO RPRO+OFFL v1.3-v1.4 CH minus CLOUDNET CH (upper left) and of the relative difference (upper right), per site. The same conventions as for **Table 14** apply. Sensing time range is indicated on the figure. Bottom panel: monthly median of FRESCO CH-CLOUDNET CH, per site. Pixel size switch (2019/08/06) and FRESCO version change (2020/11/29) are indicated on the plot.

11.3.4.2 Dispersion

Comparison with alternative S5P cloud height retrievals

The comparison of S5P CLOUD CRB CH vs S5P FRESCO CH reveals a low ½ IP68 is 0.3-0.4 km at most sites, which is below the dispersion requirement of 0.5 km (**Figure 95**, right). Exceptions are Iquique, Granada, Munich, Yacanto and Ny-Alesund where the dispersion is higher. The evolution of monthly ½ IP68 (**Figure 95**, left) does not reveal a strong variation with time, except for a decrease of the dispersion at Iquique with the upgrade of CLOUD CRB to v2. Note we have seen earlier that also the median difference CRB-FRESCO at Iquique was reduced with this upgrade.

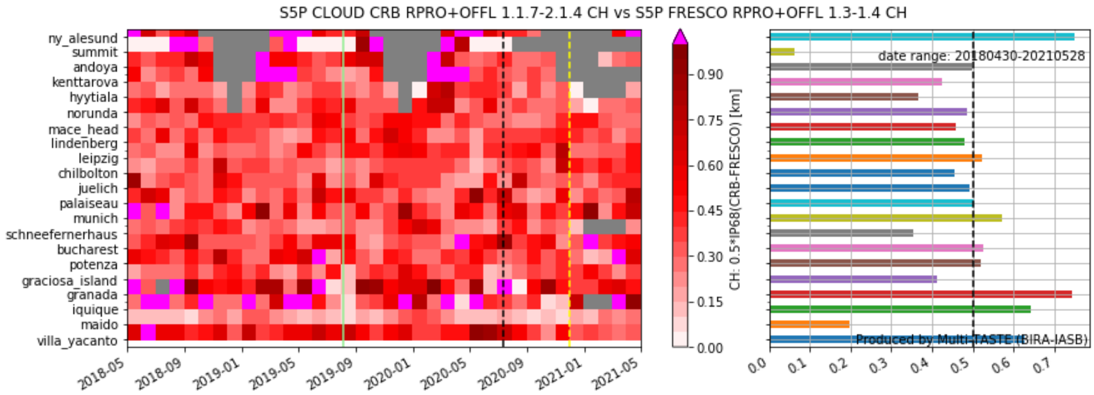


Figure 95: Left: Monthly ½ IP68 of scaled cloud fraction difference of CLOUD OFFL vs FRESCO OFFL, per month and per site. Indicated are the pixel size switch (2019/08/06), the CLOUD v1 to v2 upgrade (2020/07/13) and the FRESCO v1.3 to 1.4 upgrade (2020/11/29). Right: ½ IP68 of S5P CLOUD CRB CH minus S5P FRESCO CH. Sensing date range is indicated on the figure.

Comparison with CLOUDNET cloud top height and cloud height

From the width of the boxplots in **Figure 92** to **Figure 94**, 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 96 displays the monthly dispersion per site 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-Alesund for S5P CLOUD CAL CTH minus CLOUDNET CTH with the upgrade of CLOUD from v1 to v2.

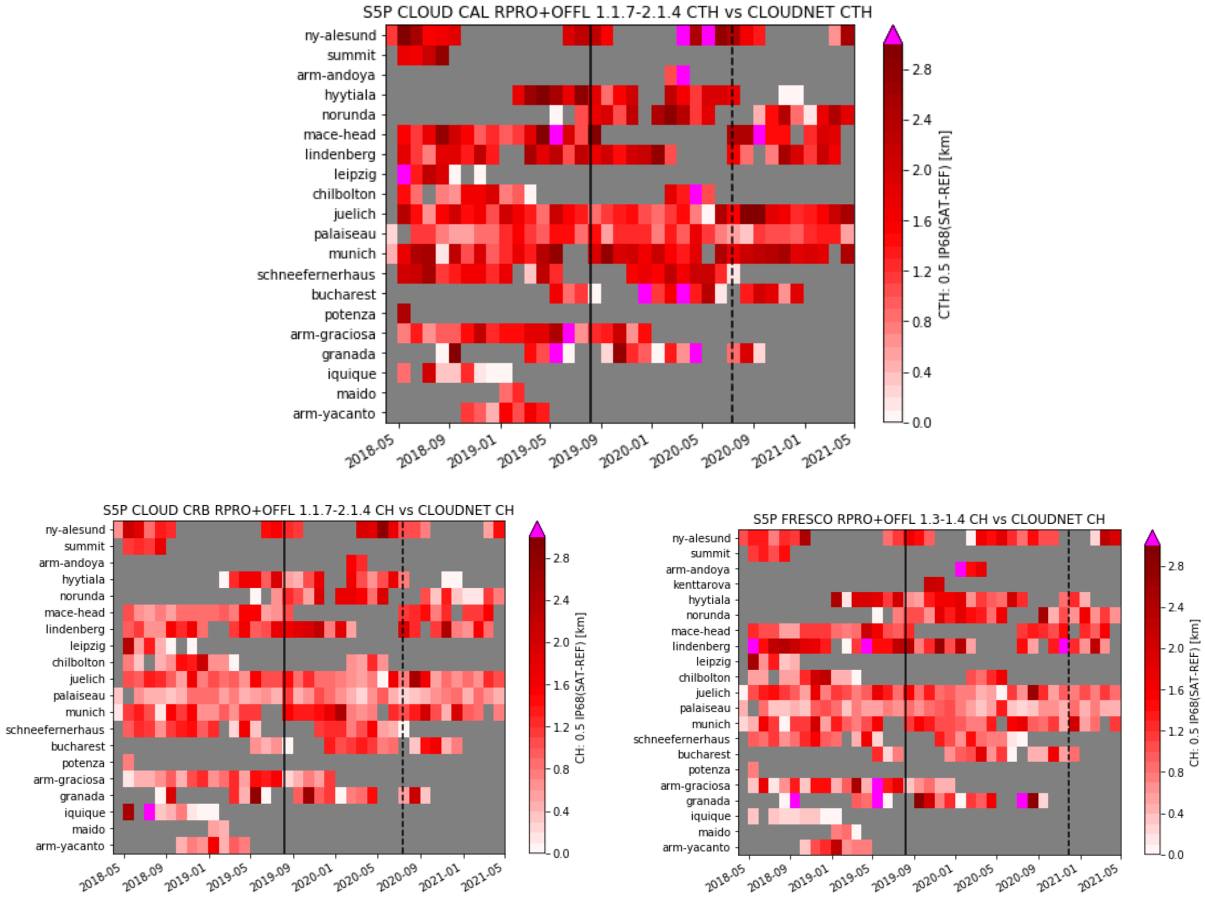


Figure 96. Monthly dispersion (½IP68) per site of S5P CLOUD CAL CTH minus CLOUDNET CTH (top), S5P CLOUD CRB CH minus CLOUDNET CH (bottom left) and S5P FRESCO CH minus CLOUDNET CH (bottom right). Indicated are the pixel size switch (2019/08/06), the CLOUD v1 to v2 upgrade (2020/07/13) and the FRESCO v1.3 to 1.4 upgrade (2020/11/29).

Overall values of dispersion are obtained by taking the median over all per-site ½ IP68 values (see **Table 14**). 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 0.5 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 ~2 km for high clouds for v1.3 and ~1.5 km for v1.4. For low clouds the dispersion is ~0.5 km.

11.3.4.3 Dependence on influence quantities

Comparison with CLOUDNET cloud top height and cloud height

Above, we have shown that at the site Iquique (a coastal site), higher discrepancies are encountered for CLOUD CRB CH – FRESCO CH (**Figure 87**, top), and CLOUD CRB CH – CLOUDNET (**Figure 93**). 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.

11.3.4.4 Short term variability

Nothing to report.

11.3.4.5 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).

11.4 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.

Figure 97 compares cloud properties of CLOUD NRTI and CLOUD OFFL at Jülich, before and after the processor version upgrade. Before the processor version upgrade, the match is about perfect. After the processor version upgrade there is still a near-perfect match for cloud fraction (not shown). For the other properties most points still match but there are also some differences.

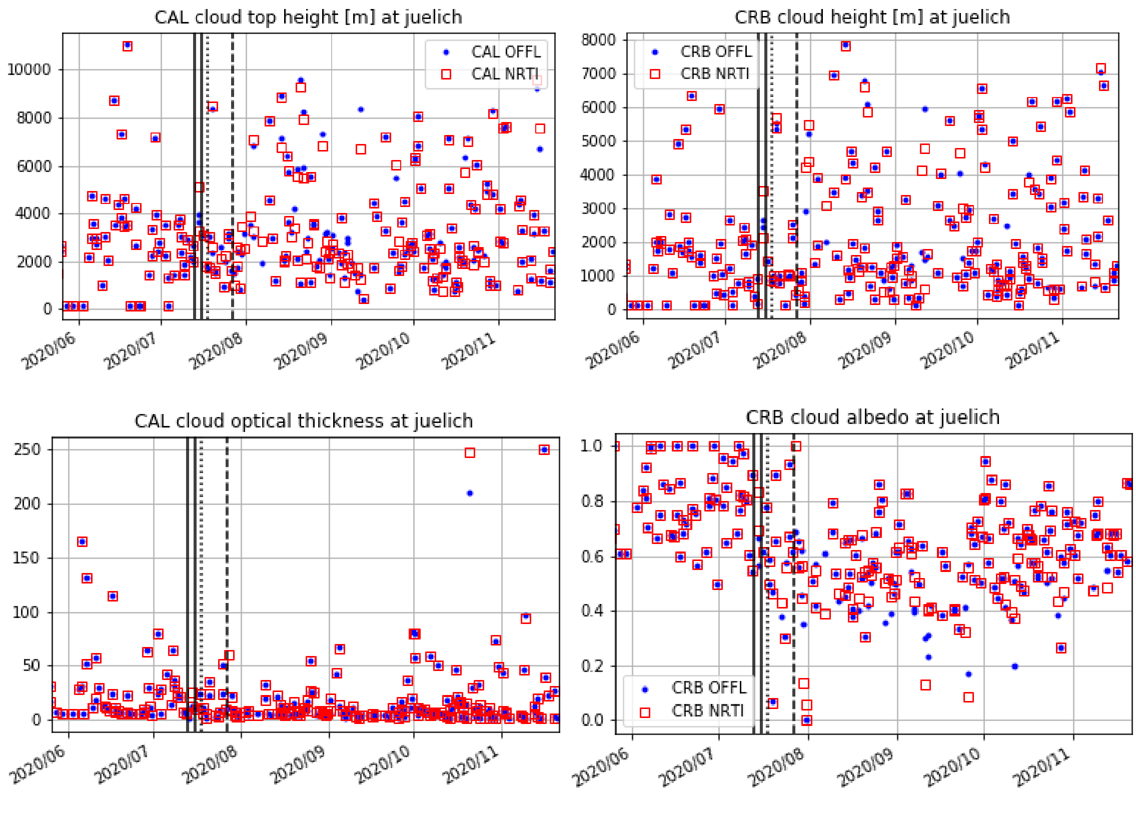


Figure 97: Time series of cloud properties of CLOUD OFFL and CLOUD NRTI at Juelich. Vertical black lines indicate the processor version change from 1.1.7 to 2.1.3; the first one for OFFL (at 2020/07/13), the second one for NRTI (at 2020/07/16).

12 Validation Results: L2_AER_AI

12.1 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 12.4 demonstrates evidence that NRTI and OFFL data do not differ significantly and that their respective validations yield similar conclusions.

12.2 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).

12.2.1 Ground-based networks

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

12.2.2 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.

12.2.3 Field campaigns and modelling support

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

12.3 Validation of L2_AER_AI NRTI

12.3.1 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].

12.3.2 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-5-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 15** 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 98**. The conclusions summarized hereafter need to be confirmed by a larger amount of test cases and co-locations, and extended over 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 15 – Case studies for different aerosol types.

Date	Type of case	TROPOMI orbit	OMI orbit	OMPS orbit
2017-11-10	Desert dust and small Sub-Saharan fire plumes	00398	70864	31285
2017-11-27	Volcanic eruption, Bali	00636	71108	31523
2017-12-13	Large biomass burning fires, California	00858	71350	31745
2018-03-31	Long-range transport of large desert dust plumes	2397, 2398	72916, 72917	33284, 33285

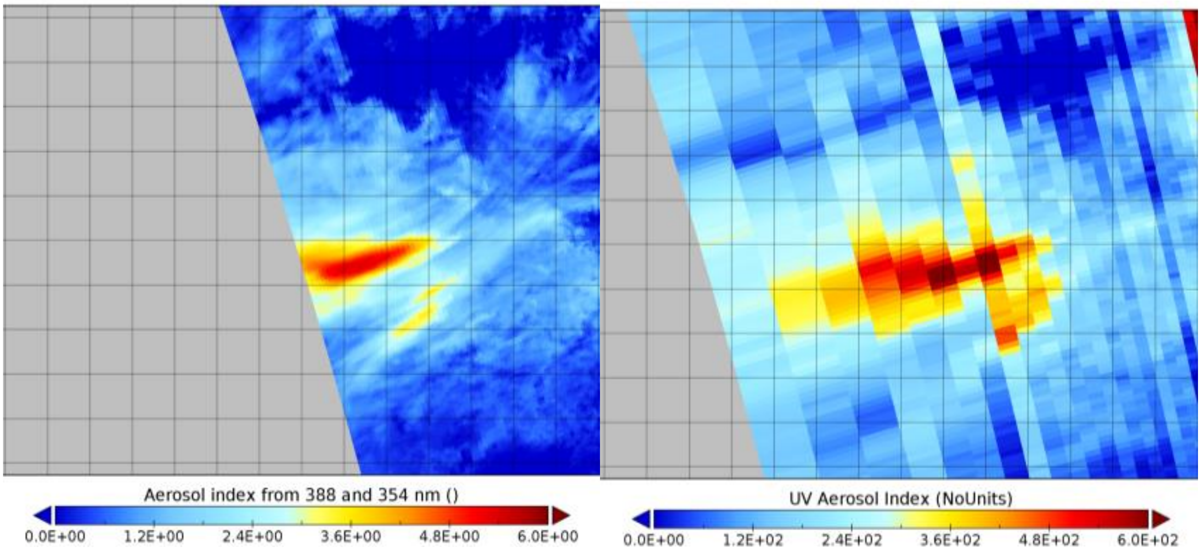
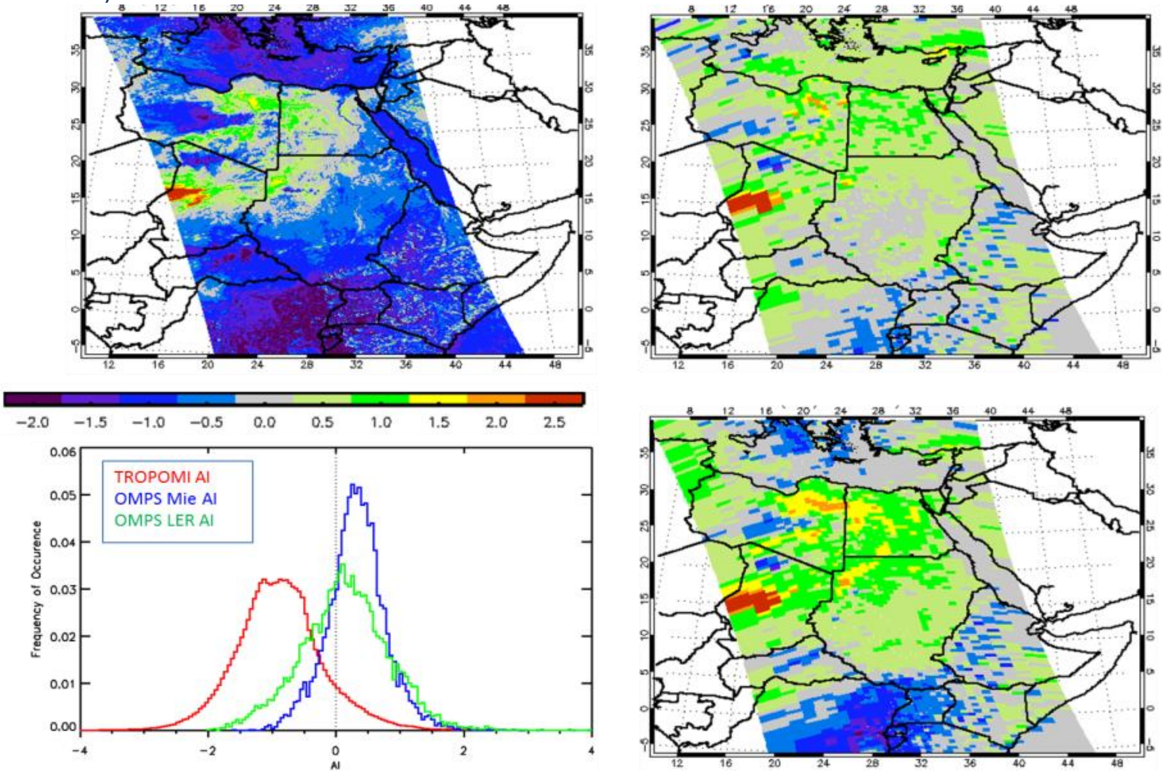


Figure 98: Comparison of S5P TROPOMI UVAI (orbit 00398, left) and OMI OMAERO UV Aerosol Index (orbit 70864, right) for Saharan dust on 10 November 2017. In general very good agreement is found (the stripes in north-south direction in the OMI data are caused by the OMI row anomaly and should be ignored).

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 99** and **Figure 100** below (courtesy of Omar Torres and Changwoo Ahn, NASA-GSFC). At the beginning of TROPOMI measurements (Nov. and Dec. 2017, **Figure 99**), the patterns of enhanced UVAI agree very well. But the S5P TROPOMI UVAI is mostly negative and is systematically smaller than the OMPS results. The negative bias of the S5P TROPOMI UVAI is steadily increasing so that is now 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 100**). 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. Also the spread of the S5P TROPOMI UVAI values has become broader than during the early phase of measurements (see **Figure 103**). Also the reason for this degradation of the data quality has to be further investigated.

Desert dust, 10 Nov 2017



Biomass burning, 12 Dec 2017

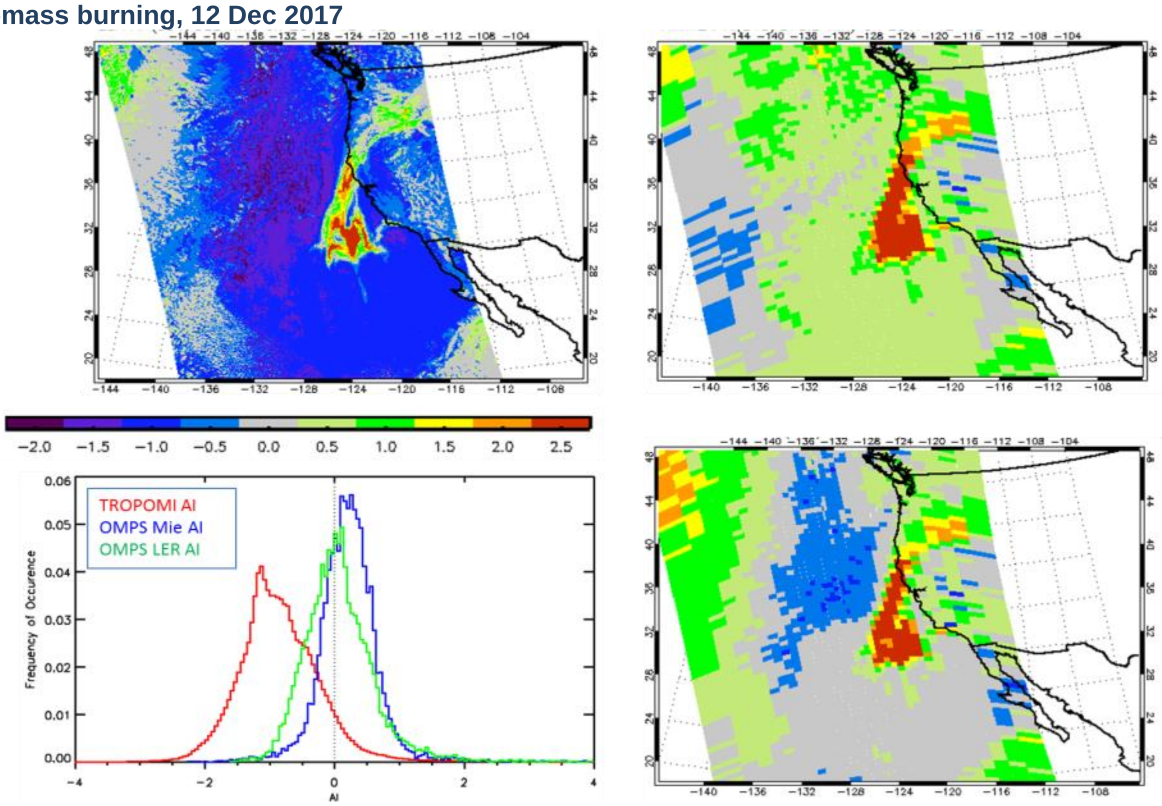


Figure 99: Comparison of UVAI from TROPOMI and OMPs for a situation with desert dust (10 Nov 2017, top) and biomass burning (12 Dec 2017, bottom). For OMPS, UVAI are calculated either assuming LER or Mie clouds. The UVAI for Mie clouds yield more consistent results. The frequency distributions indicate 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).

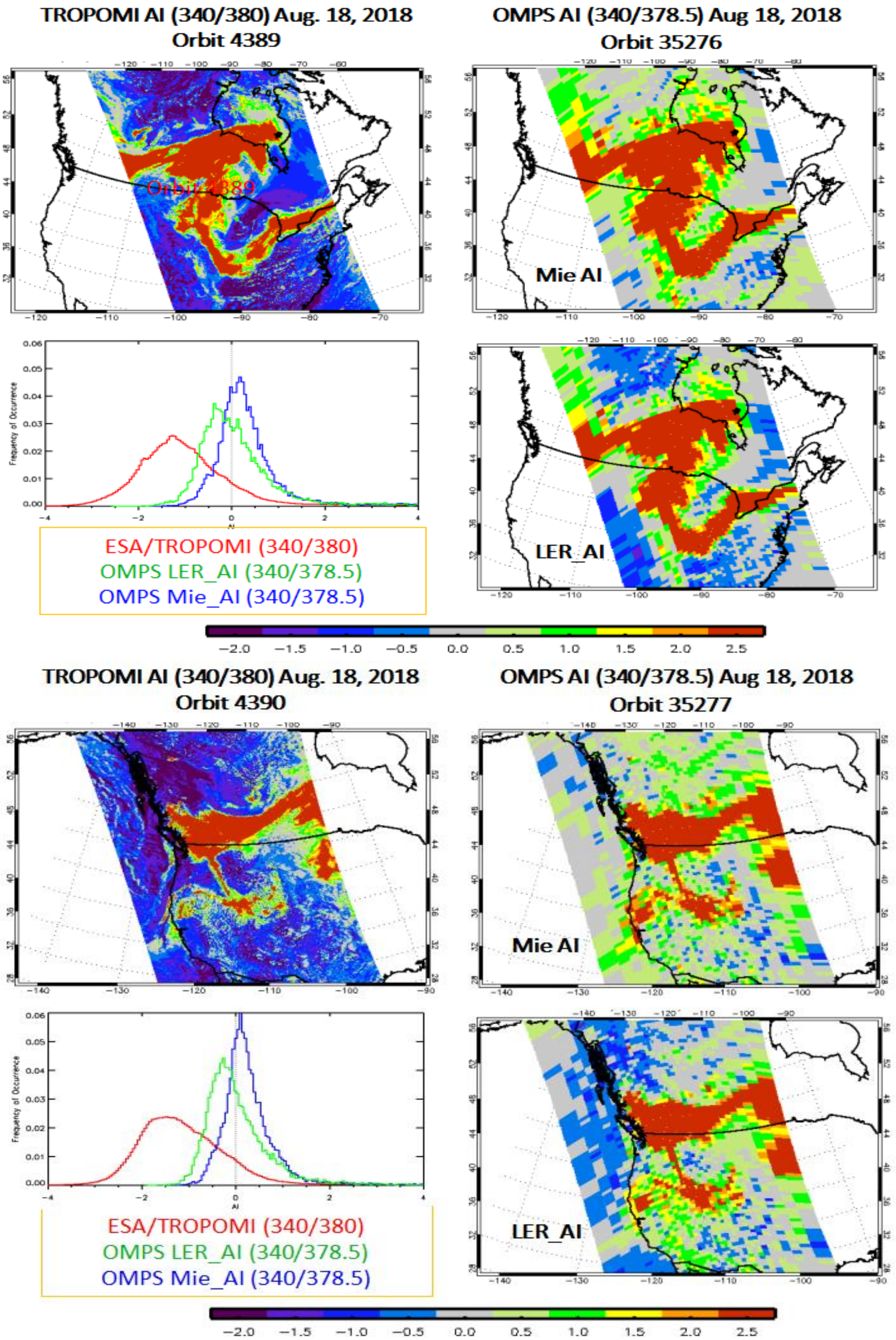


Figure 100: Comparison of UVAI from TROPOMI and OMPs for an observation of a biomass burning plume (18 Aug. 2018). For OMPS UVAI are calculated assuming either LER or Mie clouds. The UVAI for Mie clouds yields more consistent results. In contrast to early TROPOMI observations, values have systematically 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 101** below shows an example of measurements of UVAI and NO_2 VCDs, for which enhanced NO_2 and UVAI are found at the same locations.

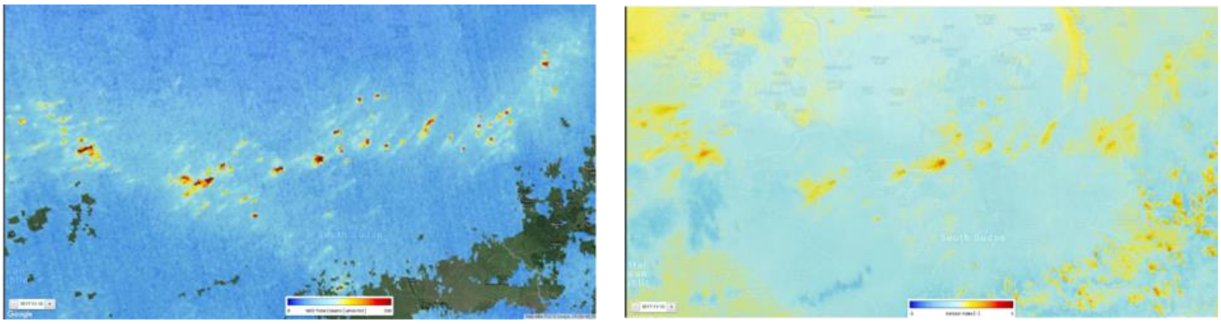


Figure 101: Comparison of NO_2 VCDs (left) and UVAI (right) obtained from S5P TROPOMI for sub-Saharan fires on 10 November 2017.

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 and will very probably be corrected after the new release of the level-1b data (foreseen in the second half of 2020). 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.

12.3.3 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 15** above conclude to a mean bias of -0.8990 AAI with OMPS (TROPOMI UVAI 354/388 – OMPS LER AI 340/378.5). Since the beginning of 2019 the UVAI is slightly outside (below 1 UVAI unit).

12.3.4 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.

12.3.5 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.

12.3.6 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 August 2020 is shown in **Figure 102** below, 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 is known feature in the L1b data and will be addressed in the next Level 1 processor update in 2020.

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.

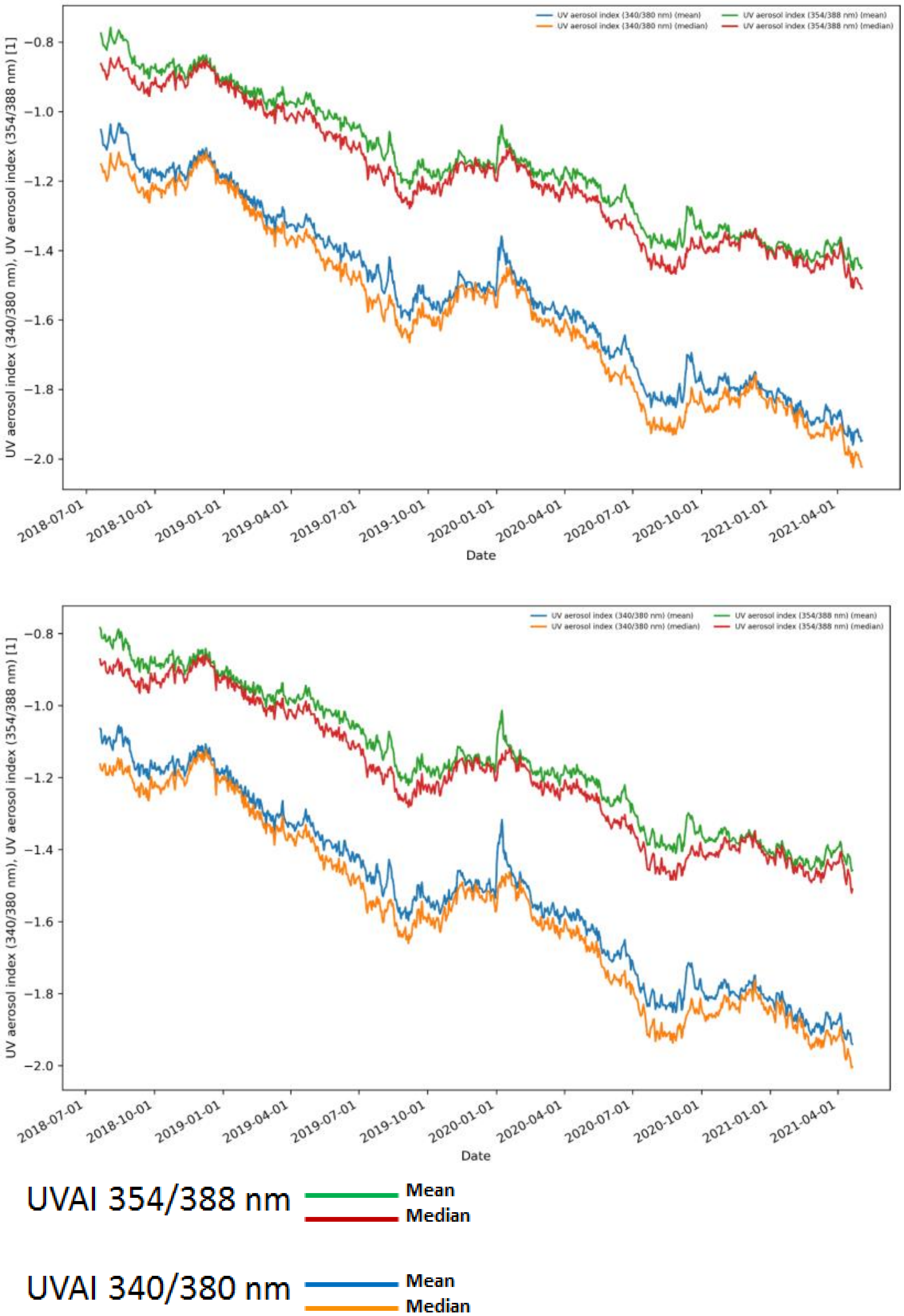


Figure 102: Comparison of the global daily mean and median for both L2_AER_AI UVAI wavelength pairs (340/380 and 354/388 nm) and for the NRTI and OFFL data streams, from 20 July 2018 through April 2021.

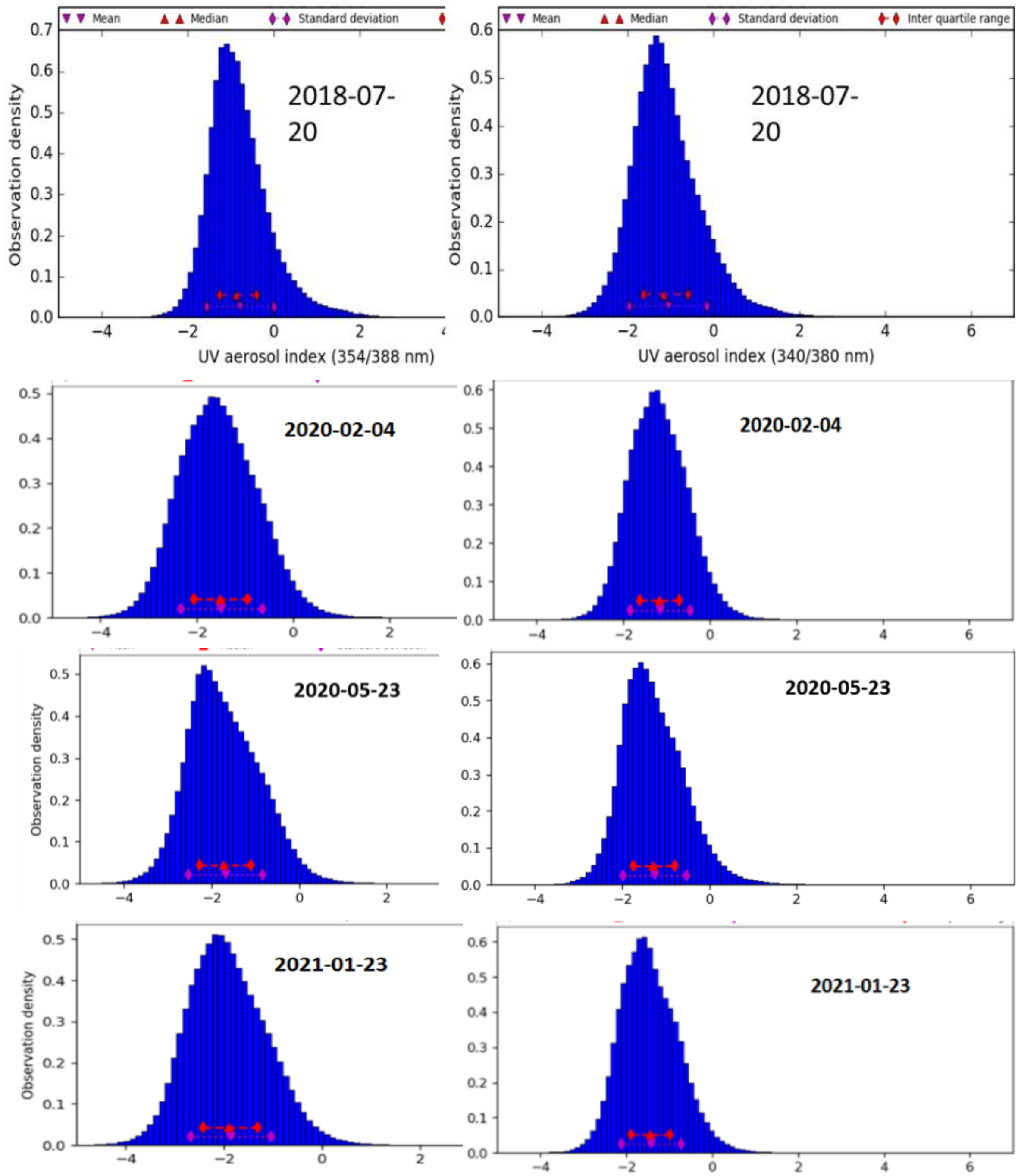


Figure 103: Comparison of the frequency distribution of the UVAI (left: 354/388m, right: 340/380nm) for four selected days (20 July 2018, 04 February 2020, 23 May 2020, and 23 January 2021).

12.3.7 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.

12.3.8 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 will be corrected once the version 2 of L1b data is made operational. Until then 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.

12.4 Equivalence of L2_AER_AI NRTI and OFFL products

Figure 104 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 also 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.

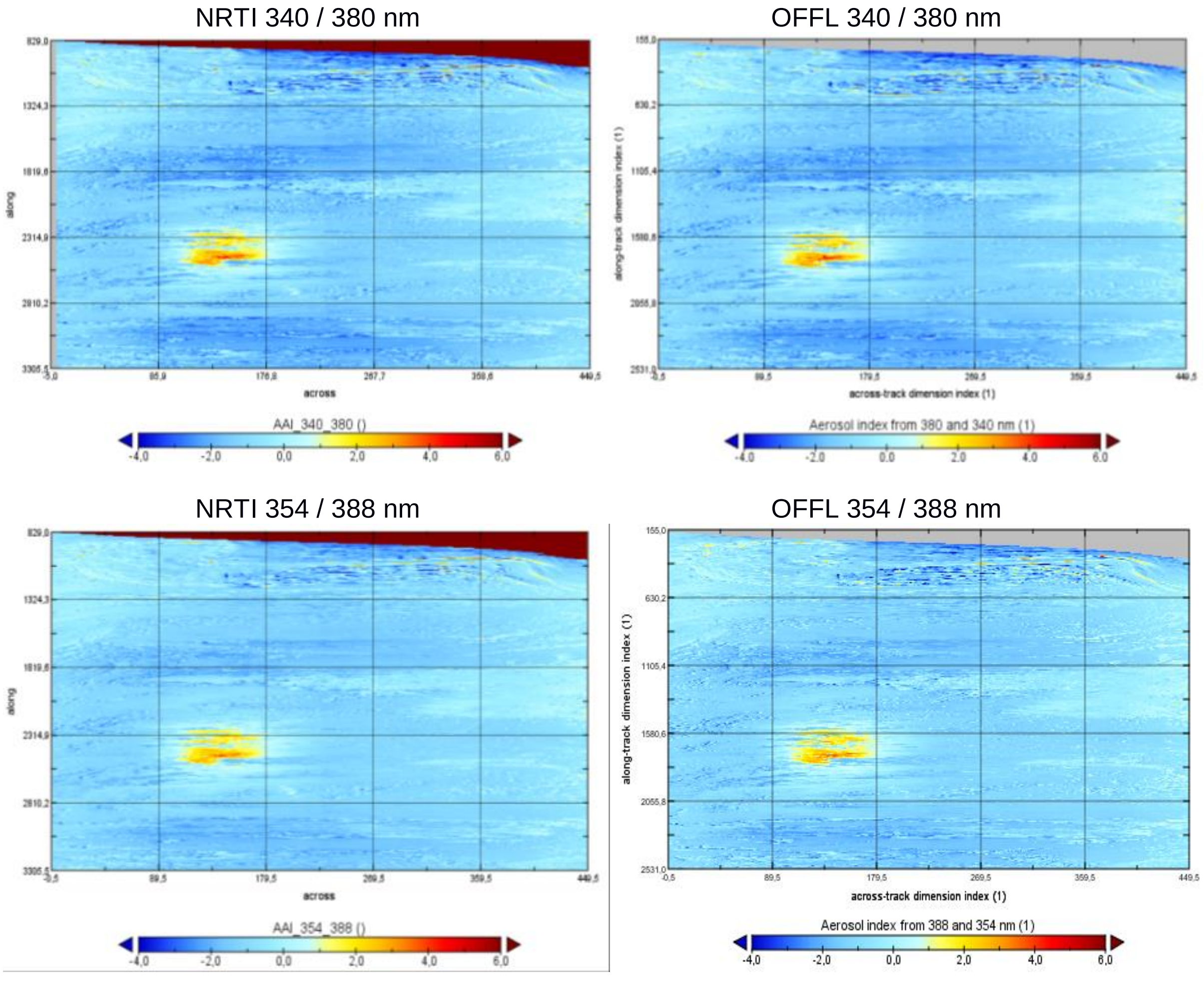


Figure 104: Comparison of the S5P TROPOMI UVAI for a selected orbit (#05033) on 3 October 2018 for the two wavelength pairs (top: 340 / 380 nm, bottom: 354 / 388 nm). While the geographical patterns are the same, the absolute values differ slightly with the NRT values (left) slightly higher than the offline values (right).

13 Validation Results: L2_AER_LH

13.1 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.

13.2 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.

13.2.1 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. However, currently, the number of ground-based observations co-located with TROPOMI during suitable aerosol events is very limited to non-existent. A small number of observations using ground-based lidars around Greece were compared to TROPOMI ALH, but this did not yet give conclusive results. The use of extensive lidar and ceilometer networks may help provide valuable validation results for the ALH.

13.2.2 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.

13.2.3 Field campaigns and modelling support

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

13.3 Validation of L2_AER_LH

13.3.1 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].

13.3.2 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 MISR and CALIOP, as detailed in 13.2.2. In **Table 16**, the details of four selected cases 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.

Table 16 – Case studies for desert dust cases with Calipso co-locations

Date	Type of Case	TROPOMI orbit	Calipso orbit start time [Day/Night]
2018-06-01	Desert Dust	3280	14:28:34 [D]
2018-06-08		3379	14:34:52 [D]
2018-06-10		3407 3408	14:22:32 [D]
2018-12-22	Smoke	7163 7174	12:55:29 [D]
2020-01-11		11640 11641	07:54:18 [N]

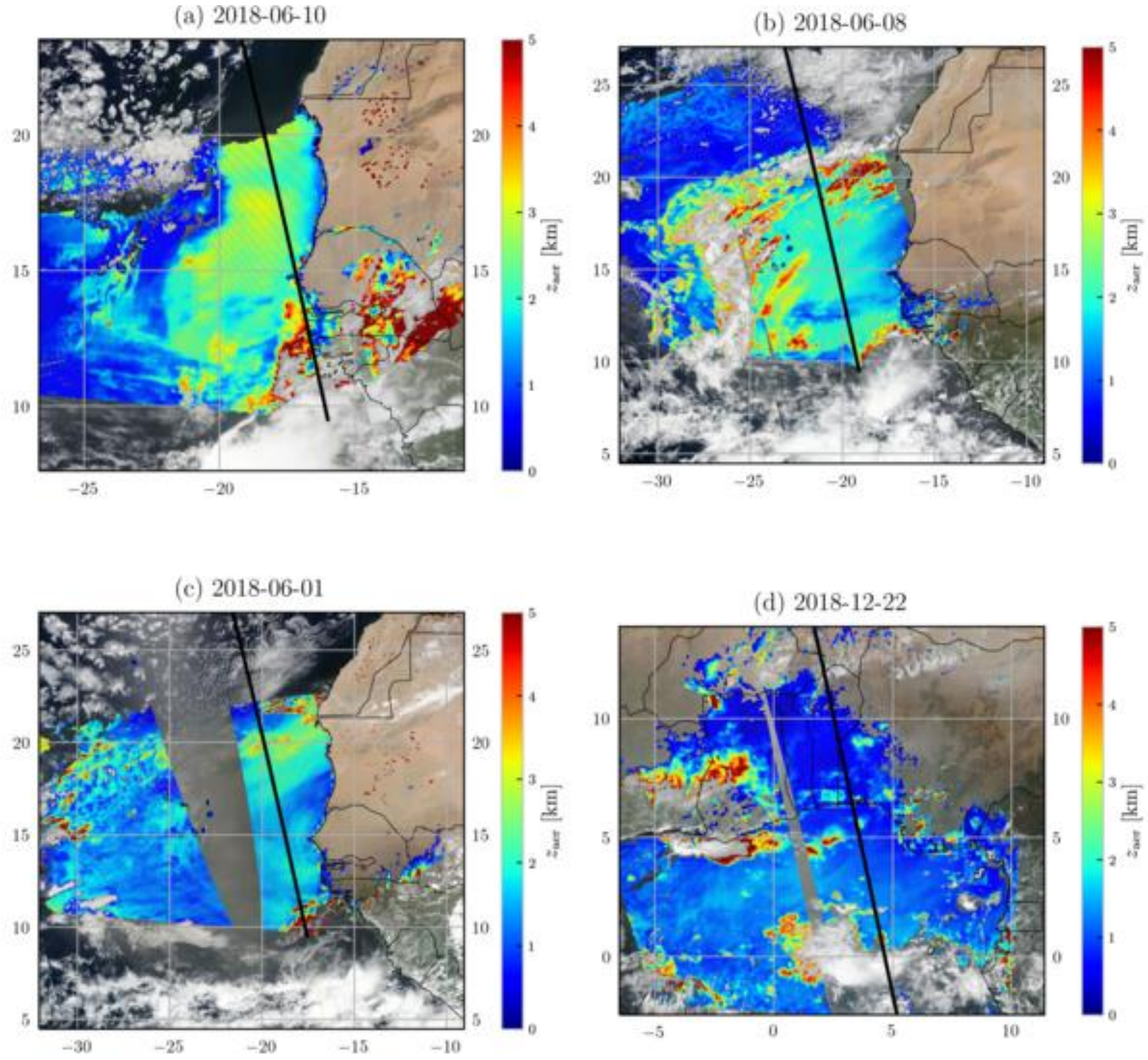


Figure 105: Details of the selected validation cases, showing the ALH on a VIIRS RGB background. The black line represents the Calipso track.

Figure 105 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 106**, 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):

$$Z_{ext} = \frac{\sum_{i=1}^{n} b_{ext,i} \cdot Z_i}{\sum_{i=1}^{n} b_{ext,i}}$$

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 106** 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.

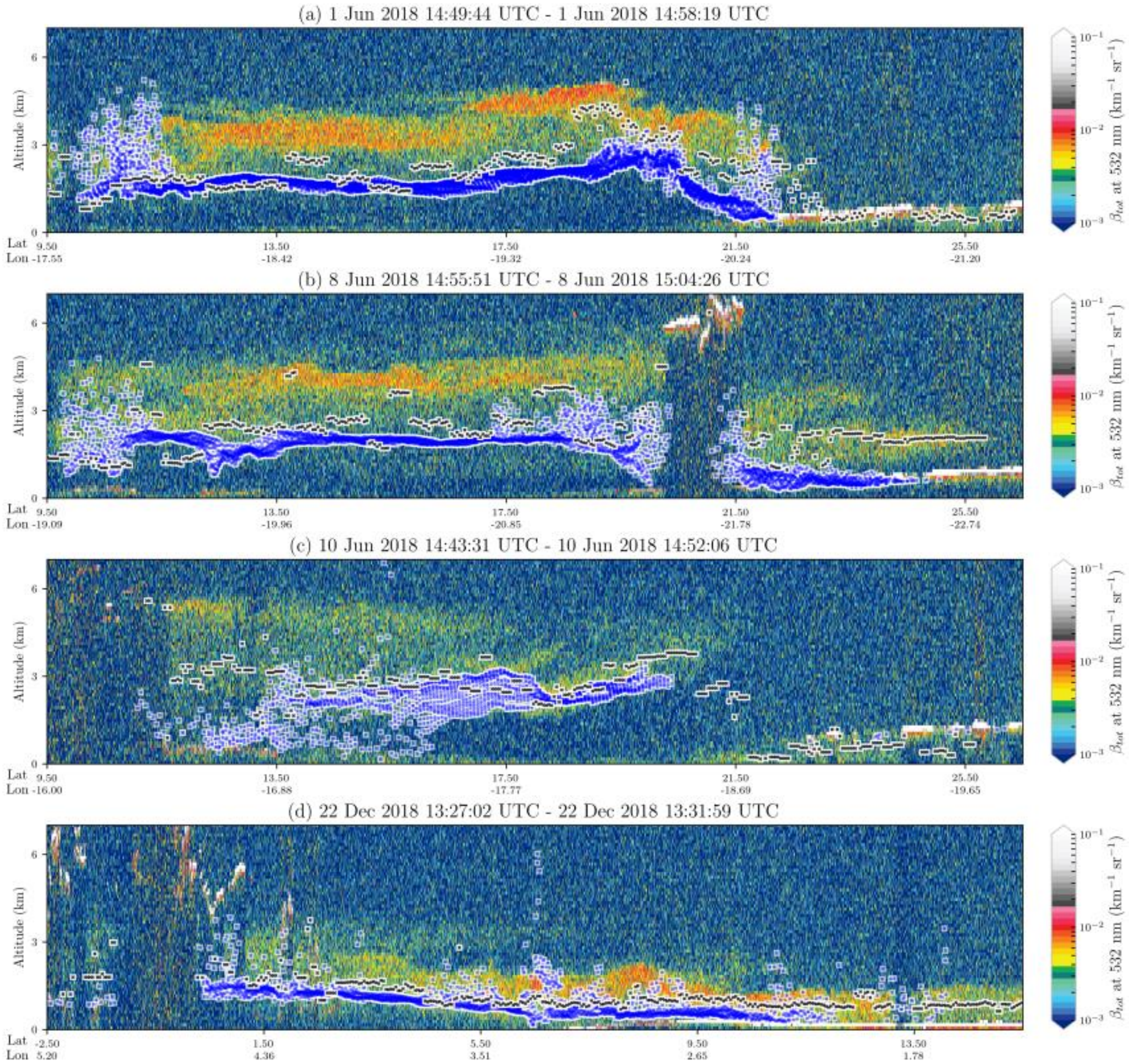


Figure 106: Curtain plots from CALIOP, showing the total attenuated backscatter at 532 nm for the four cases in **Figure 105**. Overplotted on the coloured background image of the total attenuated backscatter are the weighted extinction heights as derived from the backscatter coefficient in black-and-white, and the ALH from collocated TROPOMI pixels in blue-and-white.

In **Figure 107**, 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.

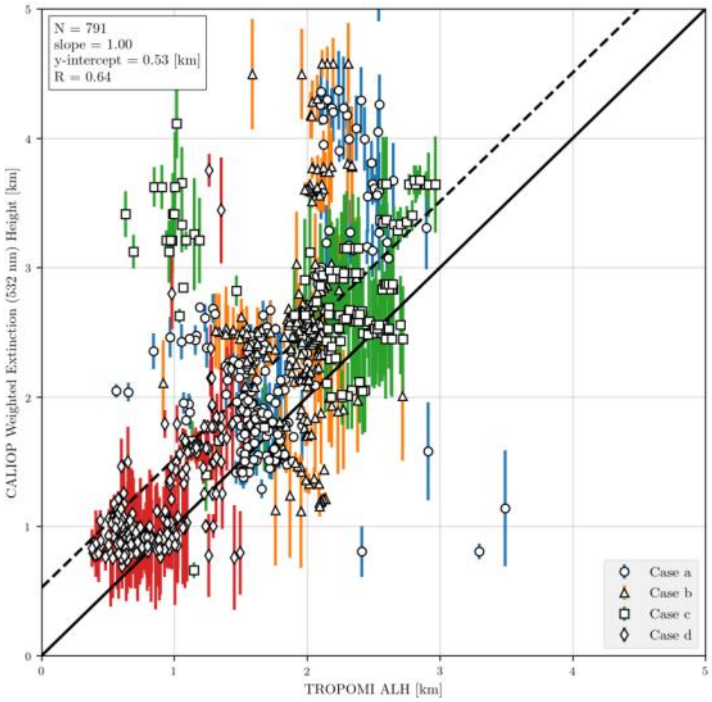


Figure 107: Comparison of ALH from TROPOMI and CALIOP for the cases presented in **Figure 105**. Each case is colour coded. CALIOP weighted extinction height is consistently lower than TROPOMI ALH.

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 108** (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 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 χ^2, smaller than 1E5. The χ^2 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 χ^2 indicate a better fit. The retrievals over land generally have much higher χ^2, and therefore are less reliable. The right panel in **Figure 108** show the results for pixels with a χ^2 than 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.

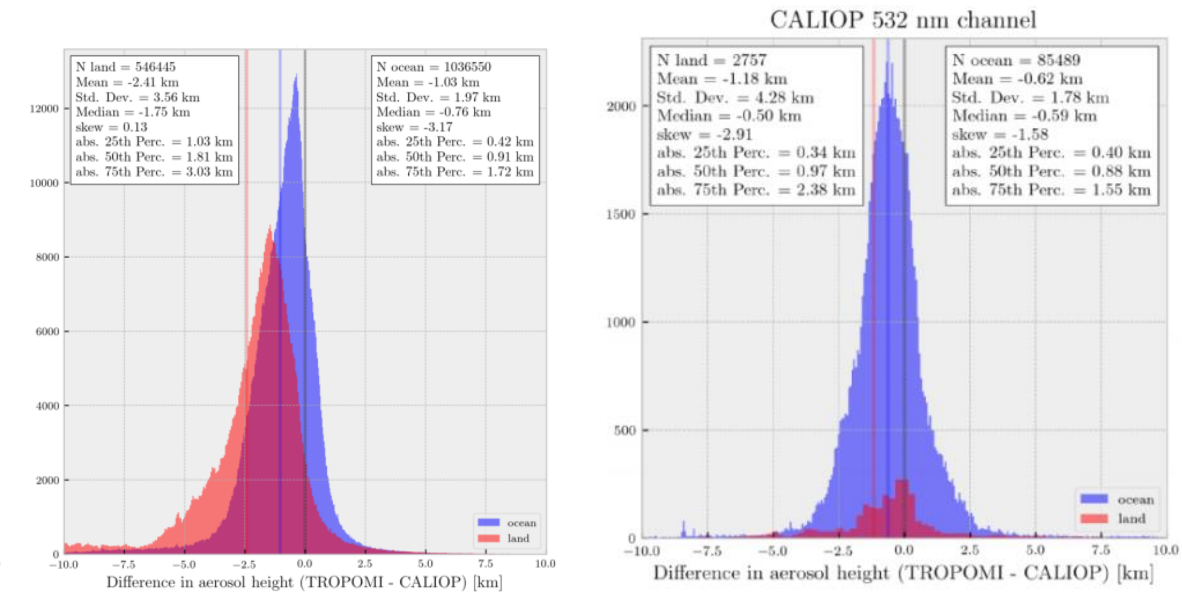


Figure 108: Histogram of differences between CALIOP weighted extinction height and TROPOMI ALH from collocated data between 1 May 2018 and 28 February 2019 (left). The right panel shows the same histogram, but for pixels that were screened for a minimal cost function (chi-squared) smaller than 1E5.

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 109** and **Figure 110**. 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.

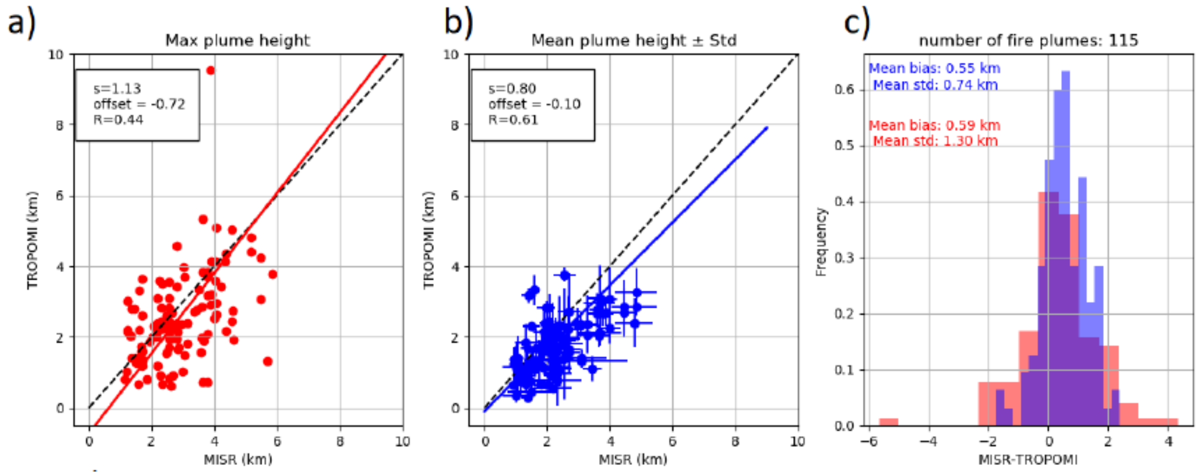


Figure 109: Comparison of TROPOMI ALH and MISR plume height for 115 fires over Nothern America in 2018. See Griffin et al, 2019 for details.

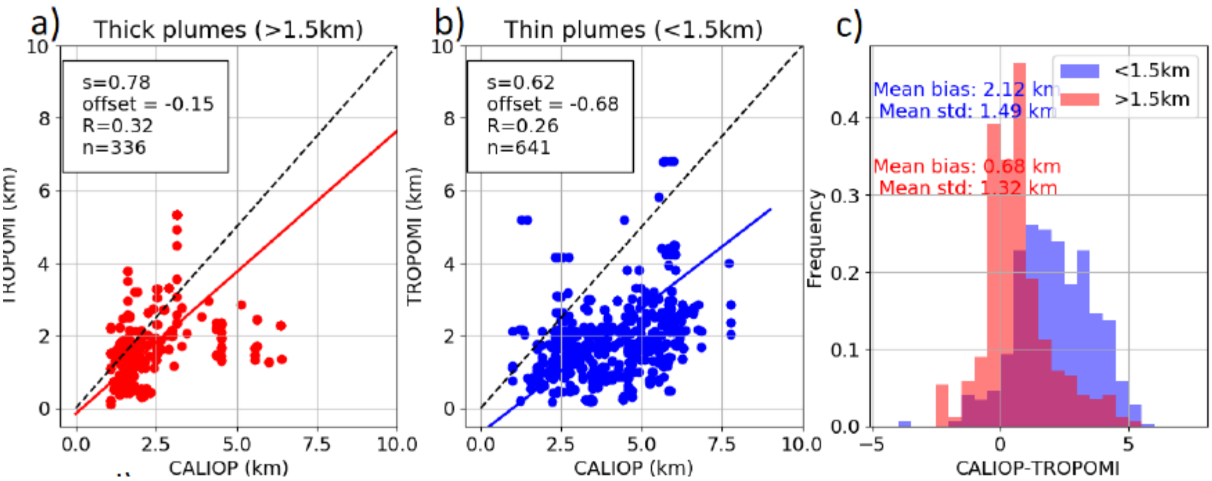


Figure 110: Comparison of TROPOMI ALH and CALIOP average aerosol layer height (top minus bottom of aerosol layer as defined by the feature mask) for collocated pixels near fires over Nothern America in 2018. See Griffin et al, 2019, for details.

13.3.3 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.

13.3.4 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. But this preliminary conclusion needs further investigation and confirmation.

13.3.5 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. The ALH is only processed for UV Aerosol Indices larger than zero. However, the UV AI is biased and degrading, which means fewer ALH pixels are processed as time continues. Currently, no significant reduction of the number of pixels can be attributed to degradation of the UV-AI. A reduction in the number of pixels is observed since August 2019 (see **Figure 111**). The figure shows the large number of retrievals during a few episodes of intense wildfires, especially around the beginning of the year 2020. This was caused by the extreme wildfires in New South Wales, producing large plumes of smoke at very high altitudes, which were visible for several weeks. Ignoring the large events, there is a gradual reduction in the number of pixels being processed visible, which is attributed to the reduction in UV-AI values. The degradation of the UV-AI will be corrected in v 2.0, to be released in Jun 2021.

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.

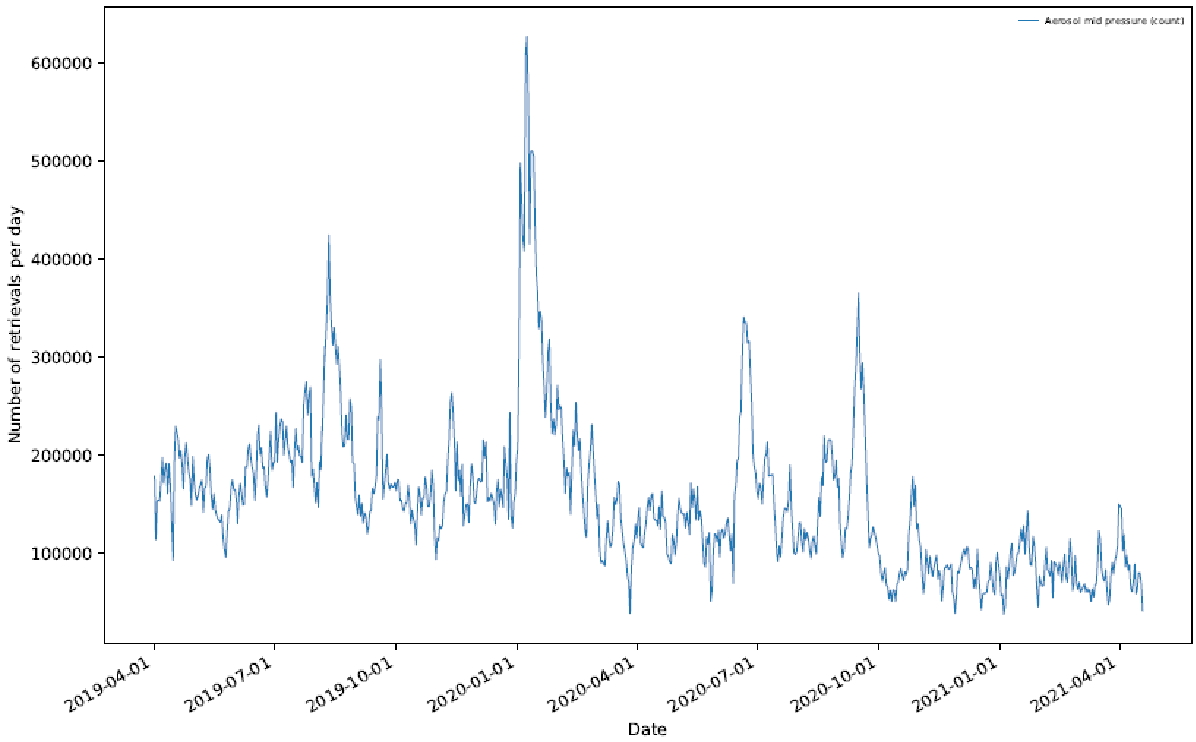


Figure 111: Number of successfully processed ALH pixels per day from April 2019 to April 2021.

13.3.6 Short term variability

The short term variability of the TROPOMI ALH was not investigated. The data product is strongly event-driven, and generally remarks on the variability currently cannot be given.

13.3.7 Geographical patterns

There are no obvious geographical features.

13.3.8 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 112** 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.

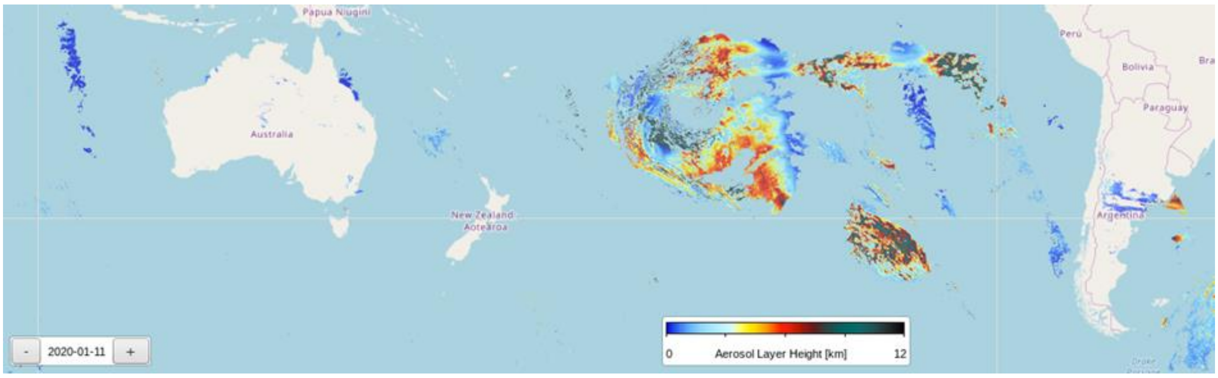


Figure 112: TROPOMI AER_LH product on 11 January 2020 over the south Pacific, showing the altitude as derived by the S5P AER_LH algorithm of the fires smoke from Australian bush fires.

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 quicklooks 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 113** 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 114** the CALIOP total attenuated backscatter at 532 nm is shown for the yellow track shown in **Figure 113**. The plume can be seen around about 44ºS and 110ºE 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.

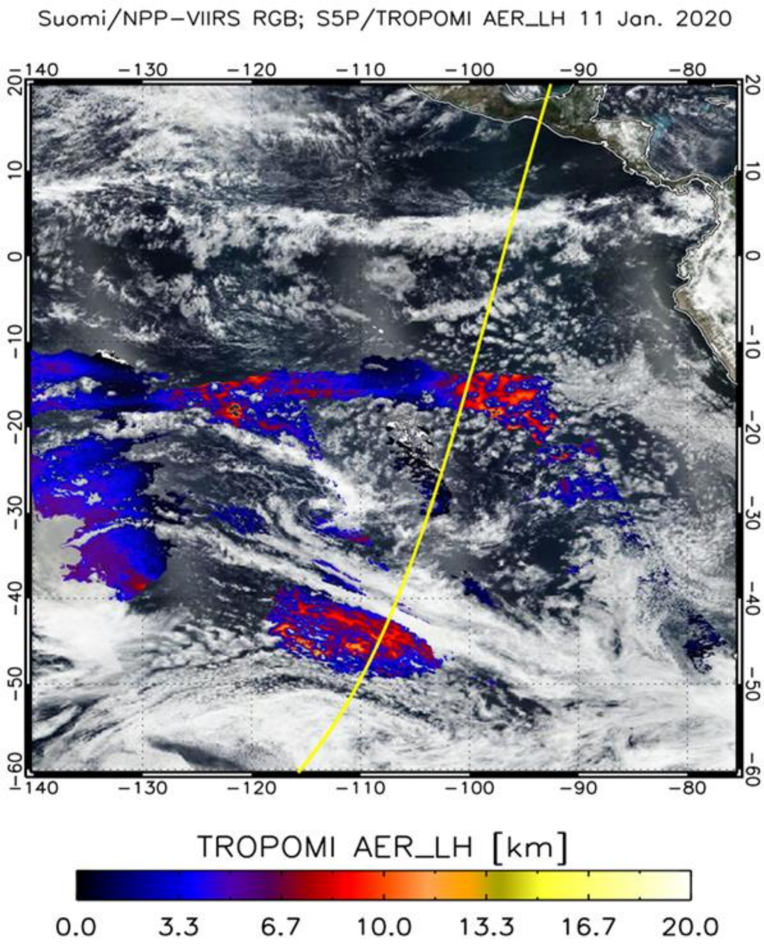


Figure 113: NPP/VIIRS RGB image with S5P/TROPOMI AER_LH on 11 January over the south Pacific with the CALIPSO track of that day overlaid in yellow.

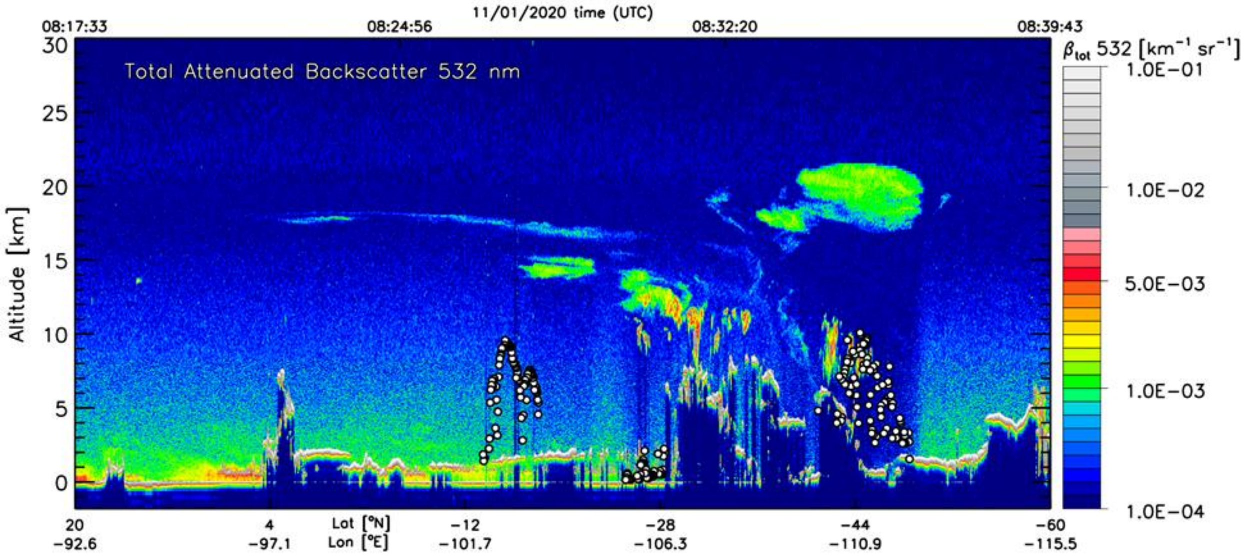


Figure 114: CALIOP aerosol backscatter rat 532 nm along the track shown in **Figure 113** (eastern of Australia on 11 January 11, 2020).

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 113** and **Figure 114**, 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.

14 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].

14.1 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

[CEOS-Nom] CEOS/ISO:19159 - Committee on Earth Observation Satellites (CEOS): general calibration and validation resources publicly available on http://calvalportal.ceos.org / ISO TS 19159-1:2014(en), Geographic information - Calibration and validation of remote sensing imagery sensors and data — Part 1: Optical sensors

[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)

[JCGM-VIM] VIM/ISO:99 Joint Committee for Guides in Metrology (JCGM/WG 2) 200:2012 & ISO/IEC Guide 99-12:2007, International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM)

[S5P-NomL1] Terms, definitions and abbreviations for TROPOMI L01b data processor;
source: KNMI; **ref:** S5P-KNMI-L01B-0004-LI; **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

14.2 Peer-reviewed articles

Basher, R. E., Review of the Dobson spectrophotometer and its accuracy, Rep. 13, WMO Global Ozone Res. and Monit. Proj., Geneva, Dec. [Available at http://www.esrl.noaa.gov/gmd/ozwv/dobson/papers/report13/report13.html], 1982.

Bauwens, M., S. Compernolle, et al.: Impact of coronavirus outbreak on NO_2 pollution assessed using TROPOMI and OMI observations, Geophys. Res. Letters, Vol. 47, Iss. 11, https://doi.org/10.1029/2020GL087978, 8 May 2020.

Bognar, K., Zhao, X., Strong, K., Boone, C., Bourassa, A., Degenstein, D., Drummond, J., Duff, A., Goutail, F., Griffin, D., Jeffery, P., Lutsch, E., Manney, G., McElroy, C., McLinden, C., Millán, L., Pazmino, A., Sioris, C., Walker, K., and Zou, J.: Updated validation of ACE and OSIRIS ozone and NO2 measurements in the Arctic using ground-based instruments at Eureka, Canada, Journal of Quantitative Spectroscopy and Radiative Transfer, 238, 106 571, https://doi.org/https://doi.org/10.1016/j.jqsrt.2019.07.014, 2019.

Compernolle, S., Argyrouli, A., Lutz, R., Sneep, M., Lambert, J.-C., Fjæraa, A. M., Hubert, D., Keppens, A., Loyola, D., O'Connor, E., Romahn, F., Stammes, P., Verhoelst, T., and Wang, P.: Validation of the Sentinel-5 Precursor TROPOMI cloud data with Cloudnet, Aura OMI O2–O2, MODIS, and Suomi-NPP VIIRS, Atmos. Meas. Tech., 14, 2451–2476, https://doi.org/10.5194/amt-14-2451-2021, 2021.

Compernolle, S., Verhoelst, T., Pinardi, G., Granville, J., Hubert, D., Keppens, A., Niemeijer, S., Rino, B., Bais, A., Beirle, S., Boersma, F., Burrows, J. P., De Smedt, I., Eskes, H., Goutail, F., Hendrick, F., Lorente, A., Pazmino, A., Piters, A., Peters, E., Pommereau, J.-P., Remmers, J., Richter, A., van Geffen, J., Van Roozendael, M., Wagner, T., and Lambert, J.-C.: Validation of Aura-OMI QA4ECV NO2 climate data records with ground-based DOAS networks: the role of measurement and comparison uncertainties, Atmos. Chem. Phys., 20, 8017–8045, https://doi.org/10.5194/acp-20-8017-2020, 2020.

De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., Boersma, F., Chan, K.-L., Donner, S., Eichmann, K.-U., Hedelt, P., Hendrick, F., Irie, H., Kumar, V., Lambert, J.-C., Langerock, B., Lerot, C., Liu, C., Loyola, D., Piters, A., Richter, A., Rivera Cárdenas, C. I., Romahn, F., Ryan, R. G., Sinha, V., Theys, N., Vlietinck, J., Wagner, T., Wang, T., Yu, H., and Van Roozendael, M.: Comparative assessment of TROPOMI and OMI formaldehyde observations against MAX-DOAS network column measurements, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2021-378, in review, 2021.

Donlon, C., and G. Zibordi, 2014, In Situ Optical Radiometry, Chapter 3 of Optical Radiometry for Ocean Climate Measurements, Eds. G. Zibordi, C. Donlon and A. Parr; Experimental Methods in the Physical Sciences Series, Vol. 47, Elsevier Inc., ISBN: 978-0-12-417011-7, 17 November 2014.

Errera, Q. and Fonteyn, D.: Four-dimensional variational chemical assimilation of CRISTA stratospheric measurements, J. Geophys. Res., 106, 12253–12265, doi:10.1029/2001JD900010, 2001.

Garane, K., Lerot, C., Coldewey-Egbers, M., Verhoelst, T., Koukouli, M. E., Zyrichidou, I., Balis, D. S., Danckaert, T., Goutail, F., Granville, J., Hubert, D., Keppens, A., Lambert, J.-C., Loyola, D., Pommereau, J.-P., Van Roozendael, M., and Zehner, C.: Quality assessment of the Ozone_cci Climate Research Data Package (release 2017) – Part 1: Ground-based validation of total ozone column data products, Atmos. Meas. Tech., 11, 1385–1402, https://doi.org/10.5194/amt-11-1385-2018, 2018.

Garane, K., Koukouli, M.-E., Verhoelst, T., Fioletov, V., Lerot, C., Heue, K.-P., Bais, A., Balis, D., Bazureau, A., Dehn, A., Goutail, F., Granville, J., Griffin, D., Hubert, D., Keppens, A., Lambert, J.-C., Loyola, D., McLinden, C., Pazmino, A., Pommereau, J.-P., Redondas, A., Romahn, F., Valks, P., Van Roozendael, M., Xu, J., Zehner, C., Zerefos, C., and Zimmer, W.: TROPOMI/S5ptotal ozone column data: global ground-based validation and consistency with other satellite missions, Atmos. Meas. Tech., https://doi.org/10.5194/amt-12-5263-2019, 2019.

Hendrick, F., Barret, B., Van Roozendael, M., Boesch, H., Butz, A., De Mazière, M., Goutail, F., Hermans, C., Lambert, J.-C., Pfeilsticker, K., and Pommereau, J.-P.: Retrieval of nitrogen dioxide stratospheric profiles from ground-based zenith-sky UV-visible observations: validation of the technique through correlative comparisons, Atmos. Chem. Phys., 4, 2091–2106, doi:10.5194/acp-4-2091-2004, 2004.

Hendrick, F., Pommereau, J.-P., Goutail, F., Evans, R. D., Ionov, D., Pazmino, A., Kyrö, E., Held, G., Eriksen, P., Dorokhov, V., Gil, M., and Van Roozendael, M.: NDACC/SAOZ UV-visible total ozone measurements: improved retrieval and comparison with correlative ground-based and satellite observations, Atmos. Chem. Phys., 11, 5975–5995, https://doi.org/10.5194/acp-11-5975-2011, 2011.

Herman, J., A. Cede, E. Spinei, G. Mount, M. Tzortziou, and N. Abuhassan, NO_2 column amounts from ground-based Pandora and MFDOAS spectrometers using the direct-sun DOAS technique: Intercomparisons and application to OMI validation, J. Geophys. Res., Vol. 114, D13307, doi:10.1029/2009JD011848, 2009.

Hubert, D., Lambert, J.-C., Verhoelst, T., Granville, J., Keppens, A., Baray, J.-L., Bourassa, A. E., Cortesi, U., Degenstein, D. A., Froidevaux, L., Godin-Beekmann, S., Hoppel, K. W., Johnson, B. J., Kyrölä, E., Leblanc, T., Lichtenberg, G., Marchand, M., McElroy, C. T., Murtagh, D., Nakane, H., Portafaix, T., Querel, R., Russell III, J. M., Salvador, J., Smit, H. G. J., Stebel, K., Steinbrecht, W., Strawbridge, K. B., Stübi, R., Swart, D. P. J., Taha, G., Tarasick, D. W., Thompson, A. M., Urban, J., van Gijsel, J. A. E., Van Malderen, R., von der Gathen, P., Walker, K. A., Wolfram, E., and Zawodny, J. M.: Ground-based assessment of the bias and long-term stability of 14 limb and occultation ozone profile data records, Atmos. Meas. Tech., 9, 2497-2534, https://doi.org/10.5194/amt-9-2497-2016, 2016.

Hubert, Daan, Klaus-Peter Heue, Jean-Christopher Lambert, Tijl Verhoelst, Marc Allaart, Steven Compernolle, Patrick D. Cullis, Angelika Dehn, Christian Félix, Bryan J. Johnson, Arno Keppens, Debra E. Kollonige, Christophe Lerot, Diego Loyola, Maznorizan Mohamad, Maria Paulete Pereira Martins, Ankie Piters, Henry Selkirk, Anne M. Thompson, Pepijn Veefkind, Holger Vömel, Jacquelyn C. Witte, and Claus Zehner, TROPOMI tropospheric ozone column data : Geophysical assessment and comparison to ozonesondes, GOME-2B and OMI, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2020-123, in review, 2020.

Inness, A., Flemming, J., Heue, K.-P., Lerot, C., Loyola, D., Ribas, R., Valks, P., van Roozendael, M., Xu, J., and Zimmer, W.: Monitoring and assimilation tests with TROPOMI data in the CAMS system: near-real-time total column ozone, Atmos. Chem. Phys., 19, 3939–3962, https://doi.org/10.5194/acp-19-3939-2019, 2019.

Keppens, A., J.-C. Lambert, J. Granville, D. Hubert, T. Verhoelst, S. Compernolle, B. Latter, B. Kerridge, R. Siddans, A. Boynard, J. Hadji-Lazaro, C. Clerbaux, C. Wespes, D.R. Hurtmans, P.-F. Coheur, J. C. A. van Peet, R. J. van der A, K. Garane, M. E. Koukouli, D. S. Balis, A. Delcloo, R. Kivi, R. Stübi, S. Godin-Beekmann, M. Van Roozendael, and C. Zehner: Quality assessment of the Ozone_cci Climate Research Data Package (release 2017): 2. Ground-based validation of nadir ozone profile data products, Atmos. Meas. Tech., 11, 3769-3800, https://doi.org/10.5194/amt-11-3769-2018, 2018.

Keppens, A., Compernolle, S., Verhoelst, T., Hubert, D., and Lambert, J.-C.: Harmonization and comparison of vertically resolved atmospheric state observations: Methods, effects, and uncertainty budget, Atmos. Meas. Tech., 12, 4379–4391, https://doi.org/10.5194/amt-12-4379-2019, 2019.

Kerr, J. B., C. T. McElroy, and R. A. Olafson, Measurements of ozone with the Brewer ozone spectrophotometer, in Proceedings of the Quadrennial Ozone Symposium, Boulder, Colorado, edited by J. London, pp. 74–79, Natl. Cent. for Atmos. Res., Boulder, Colorado, 1981.

Kerr, J. B., I. A. Asbridge, and W. F. J. Evans, Intercomparison of total ozone measured by the Brewer and Dobson spectrophotometers at Toronto, J. Geophys. Res., 93, 11,129–11,140, doi:10.1029/JD093iD09p11129, 1988.

Lambert, J.-C., M. Van Roozendael, J. Granville, P. Gerard, P. Peeters, P.C. Simon, H. Claude and J. Staehelin, Comparison of the GOME ozone and NO_2 total amounts at mid-latitude with ground-based zenith-sky measurements, in Atmospheric Ozone - 18th Quad. Ozone Symp., L'Aquila, Italy, 1996, R. Bojkov and G. Visconti (Eds.), Vol. I, pp. 301-304, 1997.

Lambert, J.-C., M. Van Roozendael, M. De Mazière, P.C. Simon, J.-P. Pommereau, F. Goutail, A. Sarkissian, and J.F. Gleason, Investigation of pole-to-pole performances of spaceborne atmospheric chemistry sensors with the NDSC, Journal of the Atmospheric Sciences, Vol. 56, pp. 176-193, https://doi.org/10.1175/1520-0469(1999)056<0176:IOPTPP>2.0.CO;2 , 1999.

Loew, A., W. Bell, L. Brocca, C. E. Bulgin, J. Burdanowitz, X. Calbet, R. V. Donner, D. Ghent, A. Gruber, T. Kaminski,J. Kinzel, C. Klepp, J.-C. Lambert, G. Schaepman-Strub, M. Schröder, and T. Verhoelst, Validation practices for satellite based earth observation data across communities, Rev. Geophys., DOI: 10.1002/2017RG000562, 2017.

Ludewig, A., Kleipool, Q., Bartstra, R., Landzaat, R., Leloux, J., Loots, E., Meijering, P., van der Plas, E., Rozemeijer, N., Vonk, F., and Veefkind, P.: In-flight calibration results of the TROPOMI payload on-board theSentinel-5 Precursor satellite, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2019-488, revised manuscript accepted for publication, 2020.

Nanda, S., Veefkind, J. P., de Graaf, M., Sneep, M., Stammes, P., de Haan, J. F., Sanders, A. F. J., Apituley, A., Tuinder, O., and Levelt, P. F.: A weighted least squares approach to retrieve aerosol layer height over bright surfaces applied to GOME-2 measurements of the oxygen A band for forest fire cases over Europe, Atmos. Meas. Tech., 11, 3263-3280, https://doi.org/10.5194/amt-11-3263-2018, 2018.

Nanda, S., de Graaf, M., Veefkind, J. P., Sneep, M., ter Linden, M., Sun, J., and Levelt, P. F.: A first comparison of TROPOMI aerosol layer height (ALH) to CALIOP data, Atmos. Meas. Tech., 13, 3043–3059, https://doi.org/10.5194/amt-13-3043-2020, 2020.

Pinardi, G., M. Van Roozendael, F. Hendrick, N. Theys, N. Abuhassan, A. Bais, A. Blechschmidt, F. Boersma, A. Cede, J. Chong, U. Friess, J. Granville, J. R. Herman, R. Holla, J. Hovila, H. Irie, Y. Kanaya, N. Kouremeti, J.-C. Lambert, J. Ma, E. Peters, A. Piters, O. Postylyakov, A. Richter, J. Remmers, M. Tiefengraber, P. Valks, T. Vlemmix, T. Wagner, and F. Wittrock: Validation of tropospheric NO2 columns measurements from GOME-2 and OMI using MAX-DOAS and direct sun network observations, submitted to Atmos. Meas. Tech., amt-2020-76, in review, 2020.

Pommereau, J. and Goutail, F.: O3 and NO2 ground-based measurements by visible spectrometry during Arctic winter and spring, Geophys. Res. Lett., 15, 891–894, https://doi.org/10.1029/GL015i008p00891, 1988.

Richter, A., M. Weber, J. P. Burrows, J.-C. Lambert, and A. van Gijsel, Validation Strategy for Satellite Observations of Tropospheric Reactive Gases, Annals of Geophysics, Vol. 56, 10.4401/AG-6335, 2013.

Roscoe, H., Johnston, P., Van Roozendael, M., Richter, A., Sarkissian, A., Roscoe, J., Preston, K., Lambert, J.-C., Hermans, C., De Cuyper, W., Dzienus, S., Winterrath, T., Burrows, J., Goutail, F., Pommereau, J.-P. D'Almeida, E., Hottier, J., Coureul, C., Ramon, D., Pundt, I., Bartlett, L., McElroy, C., Kerr, J., Elokhov, A., Giovanelli, G., Ravegnani, F., Premuda, M., Kostadinov, I., Erle, F.,Wagner, T., Pfeilsticker, K., Kenntner, M., Marquard, L., Gil, M., Puentedura, O., Yela, M., Arlander, W., Kåstad Høiskar, B., Tellefsen, C., Karlsen Tørnkvist, K., Heese, B., Jones, R., Aliwell, S., and Freshwater, R.: Slant column measurements of O3 and NO2 during the NDSC intercomparison of zenith-sky UV-visible spectrometers in June 1996, Journal of Atmospheric Chemistry, 32, 281–314, 1999.

Schneider, M., Hase, F., and Blumenstock, T.: Ground-based remote sensing of HDO/H2O ratio profiles: introduction and validation of an innovative retrieval approach, Atmos. Chem. Phys., 6, 4705-4722, https://doi.org/10.5194/acp-6-4705-2006, 2006.

Tilstra, L. G., de Graaf, M., Wang, P., and Stammes, P.: In-orbit Earth reflectance validation of TROPOMI on board the Sentinel-5 Precursor satellite, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2020-98, in review, 2020.

van Kempen, T. A., van Hees, R. M., Tol, P. J. J., Aben, I., and Hoogeveen, R. W. M.: In-flight calibration and monitoring of the Tropospheric Monitoring Instrument (TROPOMI) short-wave infrared (SWIR) module, Atmos. Meas. Tech., 12, 6827–6844, https://doi.org/10.5194/amt-12-6827-2019, 2019.

Vandaele, A. C., Fayt, C., Hendrick, F., Hermans, C., Humbled, F., Van Roozendael, M., Gil, M., Navarro, M., Puentedura, O., Yela, M., Braathen, G., Stebel, K., Tørnkvist, K., Johnston, P., Kreher, K., Goutail, F., Mieville, A., Pommereau, J.-P., Khaikine, S., Richter, A., Oetjen, H., Wittrock, F., Bugarski, S., Frieß, U., Pfeilsticker, K., Sinreich, R., Wagner, T., Corlett, G., and Leigh, R.: An intercomparison campaign of ground-based UV-visible measurements of NO2, BrO, and OClO slant columns: Methods of analysis and results for NO2, J. Geophys. Res. Atmos., 110, https://doi.org/10.1029/2004JD005423, 2005.

Verhoelst, T., Granville, J., Hendrick, F., Köhler, U., Lerot, C., Pommereau, J.-P., Redondas, A., Van Roozendael, M., and Lambert, J.-C.: Metrology of ground-based satellite validation: co-location mismatch and smoothing issues of total ozone comparisons, Atmos. Meas. Tech., 8, 5039-5062, https://doi.org/10.5194/amt-8-5039-2015, 2015.

Verhoelst, T., S. Compernolle, G. Pinardi, J.-C. Lambert, H.J. Eskes, K.-U. Eichmann, A.M. Fjæraa,J. Granville, S. Niemeijer, A. Cede, M. Tiefengraber, F. Hendrick, A. Pazmiño, A. Bais, A. Bazureau, K.F. Boersma, K. Bognar, A. Dehn, S. Donner, A. Elokhov, M. Gebetsberger, F. Goutail, M. Grutter de laMora, A. Gruzdev, M. Gratsea, G.H. Hansen, H. Irie, N. Jepsen, Y. Kanaya, D. Karagkiozidis, R. Kivi,K. Kreher, P.F. Levelt, C. Liu, M. Müller, M. Navarro Comas, A.J.M. Piters, J.-P. Pommereau, T. Portafaix, O. Puentedura, R. Querel, J. Remmers, A. Richter, J. Rimmer, C. Rivera Cárdenas, L. Saavedra de Miguel, V. P. Sinyakov, K. Strong, M. Van Roozendael, J.P. Veefkind, T. Wagner, F. Wittrock,M. Yela González, and C. Zehner, Ground-based validation of the Copernicus Sentinel-5p TROPOMI NO_2 measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks, Atmos. Meas. Tech., https://doi.org/10.5194/amt-14-481-2021, 2021.

Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, T., De Mazière, M., De Smedt, I., Grutter, M., Hannigan, J., Jones, N., Kivi, R., Lutsch, E., Mahieu, E., Makarova, M., Metzger, J.-M., Morino, I., Murata, I., Nagahama, T., Notholt, J., Ortega, I., Palm, M., Pinardi, G., Röhling, A., Smale, D., Stremme, W., Strong, K., Sussmann, R., Té, Y., van Roozendael, M., Wang, P., and Winkler, H.: TROPOMI/S5P formaldehyde validation using an extensive network of ground-based FTIR stations, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2020-30, in review, 2020.

Yela, M., Gil-Ojeda, M., Navarro-Comas, M., Gonzalez-Bartolomé, D., Puentedura, O., Funke, B., Iglesias, J., Rodríguez, S., García, O., Ochoa, H., and Deferrari, G.: Hemispheric asymmetry in stratospheric NO2 trends, Atmospheric Chemistry and Physics, 17, 13 373–389, https://doi.org/10.5194/acp-17-13373-2017, https://www.atmos-chem-phys.net/17/13373/2017/, 2017.

14.3 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_VDAF-AVS]	Validation Data Analysis Facility Automated Validation Server	http://mpc-vdaf-server.tropomi.eu
[ER_CoperATBD]	Copernicus Sentinel-5p products and algorithms documents webpage	https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms
[ER_MPS]	TROPOMI Portal for Instrument and Calibration	http://mps.tropomi.eu
[ER_L2QC]	TROPOMI Portal for Level-2 Data Quality Control	http://mpc-l2.tropomi.eu
[ER_S5PVT]	S5P Validation Team AO projects	https://earth.esa.int/web/guest/pi-community/apply-for-data/ao-s
[ER_CoperEC]	Copernicus Programme website	http://www.copernicus.eu
[ER_CoperESA]	ESA Copernicus website	http://www.esa.int/copernicus
[ER_CAMS]	Copernicus Atmosphere Monitoring Service (CAMS) website	http://atmosphere.copernicus.eu
[ER_C3S]	Copernicus Climate Change Service (C3S) website	http://climate.copernicus.eu
[ER_CEOS-Nom]	CEOS Cal/Val Terms and Definitions	http://calvalportal.ceos.org/documents/10136/551648/IASB-BIRA+Metrology+Terms+and+Definitions
[ER_GUM]	Guide to the expression of uncertainty in a measurement (GUM)	http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf
[ER_VIM]	International Vocabulary of Metrology (VIM)	http://www.bipm.org/en/publications/guides/vim.html
[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_FRM4GHG]	Fiducial Reference Measurements for Ground-Based Infrared Greenhouse Gas Observations project website	http://frm4ghg.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	http://www.actris.eu
[ER_Cloudnet]	Cloudnet remote sensing network website	http://www.cloud-net.org
[ER_EARLINET]	European Aerosol Research Lidar Network (EARLINET) website	http://www.earlinet.org
[ER_EUBREWNET]	COST Action for a coherent network of European Brewer Spectrophotometer monitoring stations (EUBREWNET) website	http://www.eubrewnet.org
[ER_EUMETNET]	European Meteorological Services Network (EUMETNET) website	http://eumetnet.eu
[ER_EVDC]	ESA Validation Data Centre (EVDC) website	http://evdc.esa.int
[ER_NDACC]	Network for the Detection of Atmospheric Composition Change (NDACC) website	http://ndacc.org
[ER_NOVAC]	Network for Observation of Volcanic and Atmospheric Change (NOVAC) website	http://novac-community.org/
[ER_SHADOZ]	Southern Hemisphere ADditional OZonesonde programme website	https://tropo.gsfc.nasa.gov/shadoz
[ER_TCCON]	Total Carbon Column Observing Network (TCCON) website	https://tccon-wiki.caltech.edu
[ER_TOLnet]	Tropospheric Ozone Lidar Network (TOLnet) website	http://www-air.larc.nasa.gov/missions/TOLNet
[ER_WOUDC]	World Ozone and Ultraviolet Data Centre (WOUDC) website	http://woudc.org

15 Acknowledgements

This Section acknowledges the authors of this report in charge of the MPC Routine Operations validation service (**Table 17**), the operators of S5P validation facilities, the providers of Fiducial Reference Measurements and other validation data, and the support provided by the Agencies.

15.1 S5P MPC Routine Operations Validation Service

Table 17 – Responsibilities for the S5P MPC 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).

S5P Product ID	Geophysical Quantity	Product Coordinator for Routine Operations Validation Activities	Product Contributors to Routine Operations Validation Activities
L1B	Radiance and irradiance	Q. Kleipool (KNMI)	A. Ludewig (KNMI)
L2_O3	O_3 total column	T. Verhoelst (BIRA-IASB)	K. Garane (AUTH) K.-P. Heue (DLR) C. Lerot (BIRA-IASB) A. Redondas (AEMET)
L2_O3_PR	O_3 profile	A. Keppens (BIRA-IASB)	O. Tuinder (KNMI)
L2_O3_TCL	O_3 tropospheric column	D. Hubert (BIRA-IASB)	K.-P. Heue (DLR)
L2_NO2	NO_2 stratospheric column	K.-U. Eichmann (IUPB)	T. Verhoelst (BIRA-IASB)
	NO_2 tropospheric column		S. Compernolle (BIRA-IASB) H. Eskes (KNMI) G. Pinardi (BIRA-IASB) P. Valks (DLR)
	NO_2 total column		
L2_SO2	SO_2 total column	T. Wagner (MPI-C)	P. Hedelt (DLR) N. Theys (BIRA-IASB)
L2_HCHO	HCHO total column	K.-U. Eichmann (IUPB)	K.L. Chan (DLR) S. Compernolle (BIRA-IASB) I. De Smedt (BIRA-IASB) C. Vigouroux (BIRA-IASB)
L2_CO	CO total column	B. Langerock (BIRA-IASB)	J. Landgraf (SRON) M.K. Sha (BIRA-IASB)
L2_CH4	CH_4 total column	M.K. Sha (BIRA-IASB)	J. Landgraf (SRON) B. Langerock (BIRA-IASB)
L2_CLOUD	Cloud Fraction	S. Compernolle (BIRA-IASB)	R. Lutz (DLR) P. Wang (KNMI)
	Cloud Height		A. Argyrouli (DLR)
	Cloud Optical Thickness		P. Wang (KNMI)
L2_AER_AI	Aerosol Absorbing Index	T. Wagner (MPI-C)	D. Stein Zweers (KNMI)
L2_AER_LH	Aerosol Layer Height		M. de Graaf (KNMI)

15.2 S5P validation facilities

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.

The MPC Validation Data Analysis Facility (VDAF) [ER_VDAF] hosted at BIRA-IASB runs the TROPOMI Automated Validation Server developed and operated jointly by s[&]t and BIRA-IASB. This server is based on the HARP toolset developed and maintained by S. Niemeijer at s[&]t.

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 additional support provided by the EC, ESA and EUMETSAT in the context of several satellite validation and metrology projects.

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

15.3 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 Ozonesonde 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 O_3 column data from the LATMOS_RT facility at IPSL/UVSQ, WMO's Ozone Mapping Centre in Thessaloniki and WOUDC in Toronto was gathered in the framework of the S5PVT AO project VALTOZ (ID #28568, PI D. Balis, AUTH, Co-Is ML. Koukouli, E. Zyrichidou, J.-C. Lambert, T. Verhoelst, J. Granville, A. Pazmiño, F. Goutail, J.-P. Pommereau, A. Bazureau, and C. Zerefos).
- Rapid delivery O_3 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 NO_2 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. Pommereau) 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 NO_2 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 Jeseck 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 CO and CH_4 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 contracted 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.

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.

15.4 Agency support

The S5P MPC 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.

16 Terms, definitions and abbreviated terms

16.1 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]
measurement 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]

16.2 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
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
COT	Cloud Optical thickness
CRB	Clouds as Reflecting Boundaries
CRG	Climate Research Group
C(T)H	Cloud (Top) Height
DLR	German Aerospace Center / 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
EUMETNET	European Meteorological Services Network
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
EVDC	ESA Atmospheric Validation Data Centre

FRM	Fiducial Reference Measurement
FTIR	Fourier Transform Infra-Red
GE_LER	Geometry-dependent Effective Lambert-Equivalent Reflectivity
GAW	Global Atmosphere Watch
GOME(-2)	Global Ozone Monitoring Experiment(-2)
GOSAT(-2)	Greenhouse gases Observing SATellite(-2)
GSFC	Goddard Space Flight Center
GUM	Guide to the Expression of Uncertainty in Measurement
IPSL/UVSQ	Institut Pierre-Simon Laplace / Université de Versailles Saint-Quentin-en-Yvelines
IUP-UB	Institute of Environmental Physics - University of Bremen
JCGM	Joint Committee for Guides in Metrology
KNMI	Koninklijk Netherlands Meteorologisch Instituut / Royal Dutch Meteorological Institute
LATMOS	Laboratoire Atmosphères, Milieux, Observations Spatiales
LER	Lambert-equivalent reflectivity
Lidar	LIght Detection And Ranging
MAX-DOAS	Multi Axis Differential Optical Absorption Spectroscopy
MetOp	polar orbiting Meteorological Operational satellite
MPC	Mission Performance Centre
MPI-C	Max Planck Institute for Chemistry
NASA	National Aeronautics and Space Administration
NDACC	Network for the Detection of Atmospheric Composition Change
NIDFORVAL	S5P NItrogen Dioxide and FORmaldehyde VALidation using NDACC and complementary FTIR and UV-Vis DOAS ground-based remote sensing data
NOVAC	Network for Observation of Volcanic and Atmospheric Change
NILU	Norsk Institutt for Luftforskning / Norwegian Institute for Air Research
NOAA	National Oceanic and Atmospheric Administration
NRT	Near Real Time
NSO	Netherlands Space Office
PANDORA	not an acronym; direct Sun UV-visible spectrometer
OFFL	Off-line
OMI	Ozone Monitoring Instrument
OMPS	Ozone Mapper and Profiling Suite
PDGS	Payload Data Ground Segment
PI	Principal Investigator
PRF	Product Readme File
PUM	Product User Manual
QA4EO	Quality Assurance framework for Earth Observation
QC	Quality Control
QWG	Quality Working Group
RAL	Rutherford Appleton Laboratory
S5P	Sentinel-5 Precursor

S5PVT	Sentinel-5 Precursor Validation Team
SAOZ	Système d'Analyse par Observation Zénithale
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CartograpHY
SHADOZ	Southern Hemisphere ADditional OZonesonde programme
SRON	Netherlands Institute for Space Research
Suomi-NPP	Suomi National Polar-orbiting Partnership
TCCON	Total Carbon Column Observing Network
TCCON4S5P	Validation of S5P Methane and Carbon Monoxide with TCCON Data
TOLNet	Tropospheric Ozone Lidar Network
TROPOMI	Tropospheric Monitoring Instrument
UVAI	Ultraviolet aerosol absorbing index
VDAF	Validation Data Analysis Facility
VIIRS	Visible Infrared Imaging Radiometer Suite
VIM	International Vocabulary of Metrology
WMO	World Meteorological Organization
WOUDC	World Ozone and Ultraviolet Data Centre
ZSL-DOAS	Zenith-Scattered-Light Differential Optical Absorption Spectroscopy

END OF DOCUMENT