



Sentinel-5 Precursor Mission Performance Centre

Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products #09: April 2018 – November 2020

Prepared by	S5P Mission Performance Centre
Reference	S5P-MPC-IASB-ROCVR-09.01.01-20201221
CI identification	DI-MPC-ROCVR
Document update	#09
Issue	09.01.01
Date of issue	2020-12-21
Status	FINAL
Distribution	PUBLIC



Science & Technology
Facilities Council



Universität Bremen

Belgian Science Policy Office



Disclaimer

The information provided in this document has been produced in the context of the Copernicus Sentinel-5 Precursor Mission Performance Centre (S5P MPC). The activities leading to those results have been contracted by the European Space Agency (ESA) and coordinated by the Royal Belgian Institute for Space Aeronomy (BIRA-IASB) under supervision of the Royal Netherlands Meteorological Institute (KNMI) as the prime contractor. All information in this document is provided “as is” and no guarantee or warranty is given that the information is fit for any particular purpose. For the avoidance of all doubts, ESA, BIRA-IASB and KNMI expressly disclaim any and all warranties, express or implied, including without limitation warranties of merchantability and fitness for a particular purpose, with respect to the information. In no event shall ESA, BIRA-IASB and KNMI be liable for any direct, indirect, incidental, punitive, or consequential damages of any kind whatsoever with respect to the information. The user thereof uses the information at its sole risk and liability.

Approval Record		
Checked by:	J.-C. Lambert (BIRA-IASB) Q. L. Kleipool (KNMI) D. Loyola (DLR) J. P. Veeffkind (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 #09: April 2018 – November 2020
Type of document	S5P MPC Routine Operations Consolidated Validation Report (ROCVR)
Document ID	S5P-MPC-IASB-ROCVR-09.01.01-20201221
ROCVR update	#09
Issue number	version 09.01.01
Date of issue	2020-12-21
Status	FINAL
Distribution	PUBLIC
Available on:	https://mpc-vdaf.tropomi.eu
Editors	J.-C. Lambert and A. Keppens (BIRA-IASB)
Contributors	S. Compernelle (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), M.L. 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 #09: April 2018 – November 2020. Lambert, J.-C., S. Compernelle, 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, M.L. 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 #09, Version 09.01.01, 173 pp., December 21, 2020.

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], a 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 both 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 routine operation data from April 2018 till November 2020. It includes validation results for the UPAS processor upgrade to version 02.01.03 operational since mid-July 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 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. 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 $\pm 2.5\%$. The instrumental switch to smaller (along-track) ground pixels on the 6th of August 2019 does not lead to any effect on 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) 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 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 +15% or +3.1 DU) and the mean dispersion of the differences (about 24% or 4.6 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. Comparisons to ozonesonde do not indicate a clear change in data quality. However, it is currently challenging to arrive at robust conclusions on a possible change in data quality due to the considerable dispersion in the difference time series and the limited v02.01.03 co-location sample.

The S5P bias exhibits seasonal patterns with amplitude exceeding mission requirements at ozonesonde stations around the Atlantic basin. During each biomass burning season the positive bias w.r.t. ozone soundings at Paramaribo (Surinam) reaches peak values of up to 10-15 DU (or 40-60%). At other Atlantic sites (e.g., Ascension Island and Natal) the seasonal pattern in the comparisons has a smaller amplitude and its temporal correlation with the burning season is smaller. 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 tropospheric, stratospheric, and total column S5P L2_NO2 data products up to version 01.03.02 (NRTI, OFFL) and 01.02.02 (RPRO) are in good overall agreement with correlative ground-based measurements. This assessment covers the period from April 2018 (start of phase E2) until November 2020. The Pandonia Global Network (PGN), the NDACC ZSL-DOAS/SAOZ, the MAX-DOAS monitoring networks, and corresponding satellite data products (OMI) are used as a reference for the validation. The detected bias depends on the amount of NO₂ in the reference measurement. Ground-based validation concludes to rather similar bias and uncertainty estimates for the L2_NO2 NRTI and OFFL/RPRO datasets. The NRTI total column are on an average 0.8% higher than the OFFL column.

The L2_NO2 stratospheric column data is compared to NDACC ZLS-DOAS UV-visible ground-based measurements at 25 stations distributed from pole to pole. Accounting for the NO₂ diurnal variation, S5P stratospheric NO₂ is generally lower by approximately 0.17 Pmolec/cm². The median bias of -5.5% is within the S5P mission requirements (0.2-0.4 Pmolec/cm²) as well as the dispersion of 0.3 Pmolec/cm², considering the combined random errors and co-location mismatches. S5P comparisons to PGN total NO₂ column data at 3 high altitude sites conclude to a bias of approximatively -10%.

The L2_NO2 tropospheric column data is compared to MAX-DOAS ground-based column data from 21 stations, where a negative bias of -36.5% is found. The bias can be positive (7%) over clean areas and as low as -54% over highly polluted areas. The bias is thus mainly within the mission requirement of 50%. The bias reduces by up to 20% when smoothing the MAX-DOAS profile with the S5P averaging kernels. The dispersion of about 3.3 Pmolec/cm² exceeds the mission precision requirements (0.7 Pmolec/cm²). Comparison of S5P tropospheric NO₂ column data to OMI measurements concludes to differences of about -20% or more in winter time over polluted regions (China, Europe) and up to +20% in the clean Pacific region.

The L2_NO2 total NO₂ column data are compared to ground-based PGN Pandora column data at 26 sites. With a strict filtering of the Pandora data, the median bias is -5.6% with a station-to-station scatter (dispersion) of 20%. Comparison results vary with the total amount of NO₂, concluding to a positive bias of 4.5% over clean sites and high mountain areas (<7 Pmolec/cm²) and a negative bias of about -24% over polluted site.

Formaldehyde

The S5P L2_HCHO tropospheric column product up to version 02.01.03 (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 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 02.01.03 in July 2020 has no impact on data quality.

The bias with respect to NDACC MAX-DOAS measurements is roughly -40%. This bias reduces to about -20% when accounting for the difference in vertical smoothing between TROPOMI and MAX-DOAS data (averaging kernels). Validation with data from 27 FTIR stations concludes to a negative bias of -29% for high emission sites (>8x10¹⁵ molec/cm²) and a positive bias of 20% for clean sites (<2.5x10¹⁵ molec/cm²) when vertical smoothing differences are minimized by application of the averaging kernels. This is broadly consistent with the MAX-DOAS results. All biases are within the mission requirements (40-80%).

The dispersion of about 9 Pmolec/cm² (5 MAX-DOAS sites) and 7-8 Pmolec/cm² (NDACC FTIR data at clean sites) is within the uncertainty mission requirements of 12 Pmolec/cm². These values are based on the use of median and median absolute deviation 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, US, and China (<20%). The dispersion of the differences is about 2 Pmolec/cm² when considering regional averaged columns.

Sulphur Dioxide

The S5P L2_SO2 (NRTI and OFFL) sulphur dioxide column data are found in general good agreement with ground-based measurements and with other satellite observations. 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 mission requirements are fulfilled. Over volcanic plumes the requirement on the bias is fulfilled, but the requirement on the random component of the uncertainty often is not fulfilled. Here it should be noted that the current random requirement is very strict (0.15 – 0.3 DU). For the often very high SO₂ column values in volcanic plumes it is unrealistic that the random requirement can strictly be fulfilled and it is recommended to reconsider this random requirement. First checks of the UPAS processor upgrade to version 02.01.03 released in July 2020 show good consistency of the SO₂ 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 similar 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 0.67 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) CRB cloud height data and CAL cloud top height data were compared with ground-based measurements from the CLOUDNET and ARM networks. For about half of the 21 stations the discrepancy remains within or narrowly exceeds the S5P data requirement on the bias (20%). However, the sensitivity of the TROPOMI NIR observations to clouds differs significantly from the sensitivity of CLOUDNET lidar/radar instruments used as a reference, and the error associated with the reference observations is also not yet included in those comparisons. For low clouds, the bias of CAL cloud top height vs CLOUDNET is within mission requirements.

The bias and dispersion between S5P L2_CLOUD CRB cloud height and CLOUDNET is broadly similar to that between S5P FRESCO and CLOUDNET, indicating that most of the discrepancy is not specific to a particular retrieval algorithm.

For S5P L2_CLOUD cloud fraction and cloud optical thickness, satellite-to-satellite intercomparisons offer better opportunities than comparisons with ground-based observations. For now the daily mean distribution of cloud fraction, cloud top height and optical thickness as a function of latitude is compared with Aqua MODIS, Aura OMI OMCLDO2, and S5P FRESCO; and similar variations are found. Furthermore, a direct comparison for multiple days with co-located and re-gridded SNPP VIIRS cloud top height and cloud optical thickness has been performed.

Two non-physical geographical patterns in cloud parameters were identified for UPAS processor version 01.0x.0x: increased values for cloud top height and cloud fraction at the right edge of the swath at certain latitudes, and a North-South gradient for cloud albedo. These patterns are reduced for processor version 02.01.03.

A preliminary investigation of the major UPAS data processor upgrade to version 02.01.03 was performed. While cloud fraction changes considerably, the scaled cloud fraction (cloud fraction*cloud albedo/0.8) is still consistent with those of S5P FRESCO and OMI OMCLDO2. A drop in cloud (top) height at the northern site Ny-Ålesund is seen and the dispersion of the difference vs CLOUDNET is reduced.

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 slightly larger than 1 UVAI unit as compared to OMI and OMPS. The reasons for this increasing bias are related to wavelength-dependent degradation and are currently 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 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.

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	2955, 2018-05-09	3943, 2018-07-18
		01.01.02	4245, 2018-08-08	5930, 2018-12-05
		01.01.05	5931, 2018-12-05	7631, 2019-04-04
		01.01.06	7631, 2019-04-04	7999, 2019-04-30
		01.01.07	7999, 2019-04-30	12482, 2020-03-11
		01.01.08	12482, 2020-03-11	14285, 2020-07-16
		02.01.03	14285, 2020-07-16	current version
	OFFL	01.01.07	7907, 2019-04-23	12431, 2020-03-07
		01.01.08	12431, 2020-03-07	14238, 2020-07-12
		02.01.03	14239, 2020-07-13	current version
L2_O3_TCL	OFFL (CCD)	01.01.07	2818, 2018-04-30	7906, 2019-04-23
		01.01.05	2824, 2018-04-30	7421, 2019-03-20
		01.01.06	7435, 2019-03-21	7791, 2019-04-15
		01.01.07	7804, 2019-04-16	12273, 2020-02-25
		01.01.08	12288, 2020-02-26	14075, 2020-07-01
L2_NO2	NRTI	02.01.03	14132, 2020-07-02	current version
		01.00.01	2955, 2018-05-09	3364, 2018-06-07
		01.00.02	3745, 2018-07-04	3946, 2018-07-18
		01.01.00	3947, 2018-07-18	5333, 2018-07-24
		01.02.00	5336, 2018-10-24	5929, 2018-12-05
		01.02.02	5931, 2018-12-05	7517, 2019-03-27
		01.03.00	7519, 2019-03-27	7999, 2019-03-30
		01.03.01	7999, 2019-03-30	9158, 2019-07-20
	RPRO	01.03.02	9159, 2019-07-20	current version
		01.02.02	2836, 2018-05-01	5235, 2018-10-17
	OFFL	01.02.00	5236, 2018-10-17	5832, 2018-11-28
		01.02.02	5840, 2018-11-29	7424, 2019-03-20
		01.03.00	7425, 2019-03-20	7906, 2019-04-23
		01.03.01	7907, 2019-04-23	8814, 2019-06-26
		01.03.02	8815, 2019-06-26	current version
L2_HCHO	NRTI	01.00.00	2955, 2018-05-09	3943, 2018-07-18
		01.01.01	3947, 2018-07-18	4244, 2018-08-08
		01.01.02	4245, 2018-08-08	5929, 2018-12-05
		01.01.05	5931, 2018-12-05	7628, 2019-04-04
		01.01.06	7636, 2019-04-04	7999, 2019-04-30
		01.01.07	7999, 2019-04-30	12482, 2020-03-11
		01.01.08	12482, 2020-03-11	14285, 2020-07-16
		02.01.03	14285, 2020-07-16	current version
	RPRO	01.01.02	2818, 2018-04-30	4147, 2018-08-01
		01.01.05	3627, 2018-06-26	5832, 2018-11-28
	OFFL	01.01.05	5832, 2018-11-27	7545, 2019-03-29
		01.01.06	7546, 2019-03-29	7906, 2019-04-23
		01.01.07	7907, 2019-04-23	12431, 2020-03-07
		01.01.08	12432, 2020-03-07	14238, 2020-07-12
		02.01.03	14239, 2020-07-13	current version
L2_SO2	NRTI	01.00.00	3202, 2018-05-27	3847, 2018-07-10
		01.01.01	3848, 2018-07-10	4146, 2018-07-31
		01.01.02	4147, 2018-07-31	5832, 2018-12-04
		01.01.07	5833, 2018-12-05	12482, 2020-03-11

Product ID	Stream	Version	In operation from (orbit #, date)	In operation until (orbit #, date)
L2_CO	OFFL	01.01.08	12482, 2020-03-11	14285, 2020-07-16
		02.01.03	14285, 2020-07-16	current version
		01.00.00	3202, 2018-05-27	3847, 2018-07-10
		01.01.01	3848, 2018-07-10	4146, 2018-07-31
		01.01.02	4147, 2018-07-31	5932, 2018-11-28
		01.01.07	5933, 2018-11-28	12482, 2020-03-11
		01.01.08	12482, 2020-03-11	14238, 2020-07-12
		02.01.03	14239, 2020-07-13	current version
	NRTI	01.02.00	5336, 2018-10-24	5929, 2018-12-05
		01.02.02	5931, 2018-12-05	7517, 2019-03-27
		01.03.00	7519, 2019-03-27	7999, 2019-04-30
		01.03.01	7999, 2019-04-30	8906, 2019-07-03
		01.03.02	8906, 2019-07-03	current version
	RPRO ²	01.02.02	5236, 2018-10-17	5346, 2018-10-25
		01.03.01	2818, 2018-04-30	5832, 2018-11-28
	OFFL	01.03.02	2463, 2018-04-04	2477, 2018-04-05
		01.02.00	5346, 2018-10-25	5832, 2018-11-28
		01.02.02	5833, 2018-11-28	7424, 2019-03-20
		01.03.00	7425, 2019-03-20	7906, 2019-04-23
		01.03.01	7907, 2019-04-23	8814, 2019-06-26
		01.03.02	8815, 2019-06-26	current version
L2_CH4	RPRO ²	01.02.02	0657, 2017-11-28	5346, 2018-10-25
		01.03.01	2818, 2018-04-30	5832, 2018-11-28
		01.03.02	2463, 2018-04-04	2477, 2018-04-05
	OFFL	01.02.02	5833, 2018-11-28	7424, 2019-03-20
		01.03.00	7425, 2019-03-20	7906, 2019-04-23
		01.03.01	7907, 2019-04-23	8814, 2019-06-26
		01.03.02	8812, 2019-06-26	current version
	NRTI	01.01.01	3947, 2018-07-18	4244, 2018-08-08
		01.01.02	4245, 2018-08-08	5929, 2018-12-05
		01.01.05	5931, 2018-12-05	7631, 2019-04-04
		01.01.06	7631, 2019-04-04	7999, 2019-04-30
		01.01.07	7999, 2019-04-30	12482, 2020-03-11
		01.01.08	12482, 2020-03-11	14285, 2020-07-16
		02.01.03	14285, 2020-07-16	current version
L2_CLOUD	OFFL	01.01.05	5833, 2018-11-28	7546, 2019-03-29
		01.01.06	7547, 2019-03-29	7906, 2019-04-23
		01.01.07	7907, 2019-04-23	12431, 2020-03-07
		01.01.08	12432, 2020-03-07	14238, 2020-07-13
		02.01.03	14239, 2020-07-13	current version
	RPRO	01.01.05	2818, 2018-04-30	5832, 2018-11-28
		01.01.07	2818, 2018-04-30	7906, 2019-04-23
	NRTI	01.02.00	5336, 2018-10-24	5929, 2018-12-05
		01.02.02	5932, 2018-12-05	7518, 2019-03-27
		01.03.00	7519, 2019-03-27	7999, 2019-04-30
		01.03.01	8000, 2019-04-30	8906, 2019-07-03
		01.03.02	8906, 2019-07-03	current version
L2_AER_AI	OFFL	01.02.00	5236, 2018-10-11	5832, 2018-11-28
		01.02.02	5833, 2018-11-28	7424, 2019-03-20
		01.03.00	7425, 2019-03-20	7906, 2019-04-23
		01.03.01	7907, 2019-04-23	8814, 2019-06-26
		01.03.02	8815, 2019-06-26	current version
	NRTI	01.03.02		current version
L2_AER_LH	OFFL	01.03.00	7425, 2019-03-20	7906, 2019-04-23
		01.03.01	7907, 2019-04-23	8814, 2019-06-26
		01.03.02	8815, 2019-06-26	current version

Product ID	Stream	Version	In operation from (orbit #, date)	In operation until (orbit #, date)
	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 and therefore processor start 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 v02.01.03 which has a dynamic surface albedo determination).
	OFFL	O ₃ column	0.3%	2%	
L2_O3_TCL	OFFL (CCD)	O ₃ tropospheric column	+15%	24%	Positive bias over biomass burning. Geographical imprints of sampling-related biases. Seasonal change of the bias.
L2_NO2	NRTI	NO ₂ troposphere	-22%	3.6 Pmolec/cm ²	Total NO ₂ bias varies with column amount: positive bias over low pollution, negative bias over high pollution (>7 Pmolec/cm ²).
		NO ₂ stratosphere	-9%	0.4 Pmolec/cm ²	
		NO ₂ total	0±50%	-	
	OFFL RPRO	NO ₂ troposphere	-22%	3.6 Pmolec/cm ²	
L2_HCHO	OFFL RPRO	NO ₂ stratosphere	-6%	0.3 Pmolec/cm ²	HCHO bias varies with column amount: positive bias over clean areas (<2 Pmolec/cm ²), negative bias over polluted areas (>8 Pmolec/cm ²).
		NO ₂ total, low	+5%	1 Pmolec/cm ²	
	NRTI	NO ₂ total, high	-24%	3 Pmolec/cm ²	
		HCHO, low	+20%	8 Pmolec/cm ²	
L2_SO2	OFFL RPRO	HCHO, high	-29%	25 Pmolec/cm ²	Lack of validation sites in areas with high SO ₂ .
	NRTI	SO ₂ column	0.2 DU	0.2 DU	
L2_CO	OFFL	SO ₂ column	0.2 DU	0.2 DU	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.
	NRTI	CO column	6.5%	5%	
L2_CH4	OFFL	CH4 column	-0.25%	0.67%	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.
L2_CLOUD	NRTI OFFL	CAL CTH (all)	-25%	2 km	CAL CTH: smaller bias and dispersion for low clouds (CLOUDNET CTH<4km) than for high clouds (CLOUDNET CTH>4km) Snow/ice conditions degrade retrievals. Occurrence of C(T)H equal to surface height at low cloud fraction. Across track CTH and CF pattern, reduced in version 02.01.03. North-South cloud albedo pattern. COT positive bias vs NPP VIIRS.
		CAL CTH (low)	-16%	0.8 km	
		CAL CTH (high)	-30%	2.1 km	
		CRB CH	-30%	1 km	
		CAL COT	+7.9 [-]	-	

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

Document Information	4
Executive Summary	5
Processing Baseline Identification	10
Representative Quality Indicators	13
Table of Contents	15
1 Introduction	20
1.1 Background information on Sentinel-5 Precursor TROPOMI	20
1.2 Mission Performance Centre – Routine Operations Validation Service	20
1.3 Purpose, scope and outline of this document	21
2 S5P Data Quality Requirements	22
3 Validation Results: L1B_RA and L1B_IR	23
3.1 L1B products	23
3.2 Recommendations for data usage followed	23
3.3 Validation approach	24
3.4 Validation of L1B NRTI	24
3.5 Validation of L1B OFFL	24
4 Validation Results: L2_O3	25
4.1 L2_O3 products and requirements	25
4.2 Validation approach	25
4.2.1 Ground-based networks	25
4.2.2 Satellites	25
4.2.3 Field campaigns and modelling support	25
4.3 Validation of L2_O3 NRTI	25
4.3.1 Recommendations for data usage followed	25
4.3.2 Status of validation	26
4.3.3 Bias	27
4.3.3.1 Switch to smaller ground pixel size and to UPAS processor v2.1.3	28
4.3.4 Dispersion	29
4.3.5 Dependence on influence quantities	29
4.3.6 Short term variability	30
4.3.7 Geographical patterns	31
4.3.8 Other features	31
4.4 Validation of L2_O3 OFFL	32
4.4.1 Recommendations for data usage followed	32
4.4.2 Status of validation	32
4.4.3 Bias	33
4.4.3.1 Switch to smaller ground pixel size and to UPAS processor v2.1.3	34
4.4.4 Dispersion	35
4.4.5 Dependence on influence quantities	35
4.4.6 Short term variability	38

4.4.7	Geographical patterns	38
4.4.8	Other features	38
5	Validation Results: L2_O3_TCL	39
5.1	L2_O3_TCL products and requirements	39
5.2	Validation approach	39
5.2.1	Ground-based networks	39
5.2.2	Satellites	39
5.2.3	Field campaigns and modelling support	39
5.3	Validation of L2_O3_TCL OFFL (CCD)	40
5.3.1	Recommendations for data usage followed	40
5.3.2	Status of validation	40
5.3.3	Bias	40
5.3.4	Dispersion	41
5.3.5	Dependence on influence quantities	41
5.3.6	Seasonal cycle and shorter term variability	41
5.3.7	Geographical patterns	42
5.3.8	Preliminary validation of UPAS processor upgrade to version 02.01.03	42
5.3.9	Other features	42
6	Validation Results: L2_NO2	46
6.1	L2_NO2 products and requirements	46
6.2	Validation approach	46
6.2.1	Ground-based monitoring networks	46
6.2.2	Satellites	48
6.2.3	Field campaigns and modelling support	48
6.3	Validation of L2_NO2	48
6.3.1	Recommendations for data usage	48
6.3.2	Status of validation	48
6.3.3	Stratospheric NO ₂ column	49
6.3.4	Tropospheric NO ₂ column	54
6.3.5	Total NO ₂ column	62
6.4	Equivalence of L2_NO2 NRTI and OFFL products	65
6.4.1	Stratospheric NO ₂ Column NRTI vs OFFL	65
6.4.2	Tropospheric NO ₂ Column NRTI vs OFFL	66
6.4.3	Total NO ₂ Column NRTI vs OFFL	68
6.5	Internal consistency of the validation results	69
6.5.1	NO ₂ cross-sections:	69
6.5.2	ZSL-DOAS and mountain-top PGN:	70
6.5.3	Sites with multiple instruments (different geometries):	70
7	Validation Results: L2_HCHO	71
7.1	L2_HCHO products and requirements	71
7.2	Validation approach	71
7.2.1	Ground-based monitoring networks	71

7.2.2	Satellites	72
7.2.3	Field campaigns and modelling support	72
7.3	Validation of L2_ HCHO	72
7.3.1	Recommendations for data usage	72
7.3.2	Status of validation	72
7.3.3	Bias	73
7.3.4	Dispersion	77
7.3.5	Dependence on influence quantities	78
7.3.6	Short term variability	78
7.3.7	Geographical patterns	81
7.3.8	Preliminary validation of UPAS processor upgrade to version 02.01.03	81
7.4	Equivalence of L2_ HCHO NRTI and OFFL products	83
7.4.1	Bias	83
7.4.2	Dispersion	83
8	Validation Results: L2_ SO2	85
8.1	L2_ SO2 products and requirements	85
8.2	Validation approach	85
8.2.1	Ground-based networks	85
8.2.2	Satellites	85
8.2.3	Field campaigns and modelling support	85
8.2.4	Test of the expectation of zero SO ₂ SCDs (within detection limit) outside volcanic plumes and strongly polluted regions	85
8.3	Validation of L2_ SO2 NRTI	86
8.3.1	Recommendations for data usage followed	86
8.3.2	Status of validation	86
8.3.3	Bias	90
8.3.4	Dispersion	90
8.3.5	Dependence on influence quantities	91
8.3.6	Short term variability	91
8.3.7	Geographical patterns	91
8.3.8	Other features	91
8.4	Equivalence of L2_ SO2 NRTI and OFFL products	92
9	Validation Results: L2_ CO	93
9.1	L2_ CO products and requirements	93
9.2	Validation approach	93
9.2.1	Ground-based networks	93
9.2.2	Satellites	93
9.2.3	Field campaigns and modelling support	93
9.3	Validation of L2_ CO NRTI	94
9.3.1	Recommendations for data usage followed	94
9.3.2	Status of validation	94
9.3.3	Bias	95

9.3.4	Dispersion	97
9.3.5	Dependence on influence quantities	98
9.3.6	Short term variability	99
9.3.7	Geographical patterns	100
9.3.8	Other features	101
9.4	Equivalence of L2_CO OFFL and NRTI products	102
10	Validation Results: L2_CH4	103
10.1	L2_CH4 products and requirements	103
10.2	Validation approach	103
10.2.1	Ground-based networks	103
10.2.2	Satellites	103
10.2.3	Field campaigns and modelling support	103
10.3	Validation of L2_CH4 OFFL	103
10.3.1	Recommendations for data usage followed	103
10.3.2	Status of validation	104
10.3.3	Bias	105
10.3.4	Dispersion	108
10.3.5	Dependence on influence quantities	108
10.3.6	Short term variability	109
10.3.7	Geographical patterns	110
10.3.8	Change in Cloud Mask	110
10.3.9	Other features	111
11	Validation Results: L2_CLOUD	112
11.1	L2_CLOUD products and requirements	112
11.2	Validation approach	112
11.2.1	Ground-based networks	112
11.2.2	Satellites	113
11.2.3	Alternative S5P cloud algorithms	114
11.2.4	Field campaigns and modelling support	114
11.3	Validation of L2_CLOUD OFFL	115
11.3.1	Recommendations for data usage followed	115
11.3.2	Status of validation	115
11.3.3	Radiometric cloud fraction (L2_CLOUD CAL & L2_CLOUD CRB)	116
11.3.4	Cloud top height (L2_CLOUD CAL) and cloud height (L2_CLOUD CRB)	120
11.3.5	Cloud optical thickness (L2_CLOUD CAL)	134
11.3.6	Cloud albedo (L2_CLOUD CRB)	134
11.4	Preliminary validation of UPAS processor upgrade to version 2.1.3	134
11.5	Comparison of L2_CLOUD NRTI and OFFL products	137
12	Validation Results: L2_AER_AI	138
12.1	L2_AER_AI products and requirements	138
12.2	Validation approach	138
12.2.1	Ground-based networks	138

12.2.2	Satellites	138
12.2.3	Field campaigns and modelling support	138
12.3	Validation of L2_AER_AI NRTI.....	139
12.3.1	Recommendations for data usage followed	139
12.3.2	Status of validation	139
12.3.3	Bias	143
12.3.4	Dispersion	143
12.3.5	Dependence on influence quantities	143
12.3.6	Short term variability	143
12.3.7	Geographical patterns	145
12.3.8	Other features	146
12.4	Equivalence of L2_AER_AI NRTI and OFFL products.....	147
13	Validation Results: L2_AER_LH	148
13.1	L2_AER_LH products and requirements.....	148
13.2	Validation approach	148
13.2.1	Ground-based networks	148
13.2.2	Satellites	149
13.2.3	Field campaigns and modelling support	149
13.3	Validation of L2_AER_LH	149
13.3.1	Recommendations for data usage followed	149
13.3.2	Status of validation	150
13.3.3	Bias	155
13.3.4	Dispersion	156
13.3.5	Dependence on influence quantities	156
13.3.6	Short term variability	157
13.3.7	Geographical patterns	157
13.3.8	Other features	157
14	References	160
14.1	Reference documents.....	160
14.2	Peer-reviewed articles	160
14.3	Electronic references	165
15	Acknowledgements.....	167
15.1	S5P MPC Routine Operations Validation Service	167
15.2	S5P validation facilities	168
15.3	Validation data	168
15.4	Agency support	169
16	Terms, definitions and abbreviated terms	170
16.1	Terms and definitions	170
16.2	Acronyms and abbreviations	171

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 km² before and 5.5 x 3.5 km² after the operations switch to smaller ground pixel size activated on the 6th of August 2019, is accomplished thanks to a Sun-synchronous polar orbit (equator crossing time of 13:30 local solar time) and a wide swath width of 2600 km across track. From the TROPOMI radiometric measurements of Earth's radiance and solar irradiance, on-ground data processors retrieve the atmospheric abundance of ozone (O₃), nitrogen dioxide (NO₂), formaldehyde (HCHO), sulphur dioxide (SO₂), carbon monoxide (CO), methane (CH₄), as well as cloud and aerosol properties.

The S5P mission is a key component of the space segment of the European Earth Observation programme Copernicus [ER_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 12**) 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 #09 of the ROCVR presents quality information for the S5P operational data products obtained in nominal mode from April 2018 until November 2020. For data products generated with the UPAS processor (O₃, HCHO, SO₂, clouds) it includes validation results for the processor upgrade to version 02.01.03 activated on July 16, 2020. This document is structured as follows:

Document Information	4
Executive Summary	5
Processing Baseline Identification	10
Representative Quality Indicators	13
Table of Contents	15
1 Introduction	20
2 S5P Data Quality Requirements	22
3 Validation Results: L1B_RA and L1B_IR	23
4 Validation Results: L2_O3	25
5 Validation Results: L2_O3_TCL	39
6 Validation Results: L2_NO2	46
7 Validation Results: L2_HCHO	71
8 Validation Results: L2_SO2	85
9 Validation Results: L2_CO	93
10 Validation Results: L2_CH4	103
11 Validation Results: L2_CLOUD	112
12 Validation Results: L2_AER_AI	138
13 Validation Results: L2_AER_LH	148
14 References	160
15 Acknowledgements	167
16 Terms, definitions and abbreviated terms	170

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). Maintenance of these requirements is supported by the Sentinel-5p Quality Working Group (QWG), who agreed e.g. to adopt for tropospheric ozone column data the requirements expressed by the Climate Research Group (CRG) within ESA's cci_ozone project, and also to adopt maximum values of the estimates instead of ranges. 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%	1.6%-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	Enhanced total SO ₂	total column	30%	0.3 (0.12) DU
	Total SO ₂	total column	50%	1-3 (1.2) DU
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]).

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. Initial validation of the UPAS processor upgrade to version 02.01.03 – on 16 July 2020 for NRTI and on 13 July 2020 for OFFL – is also reported.

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 280];
- `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>, in <https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5-precursor-workshop-2019/sentinel-5p>, and in <http://s5pvt.skytek.com/group/s5pvt-qwg/group-documents>, 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 November 2020, and on the reference data available at the time of this report: typically, until end of October 2020 for the Dobson and Brewer data, and up to the end of November 2020 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 1**).

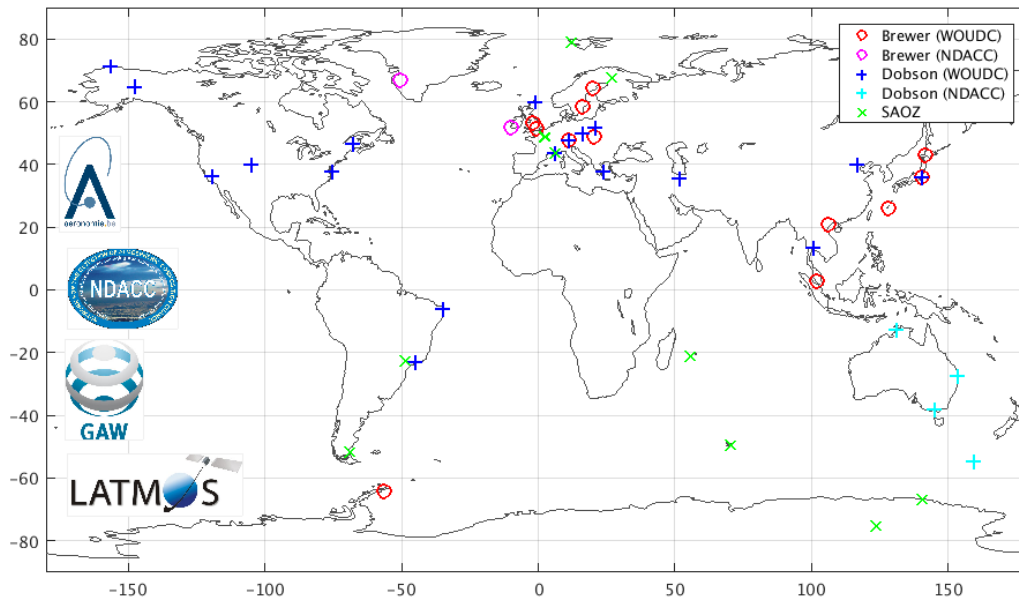


Figure 1: Geographical distribution of Brewer, Dobson and ZSL-DOAS ground-based stations for which suitable co-locations with S5P L2_O3 NRTI ozone data have been identified (May 2018 until November 2020).

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 3.5-5%).

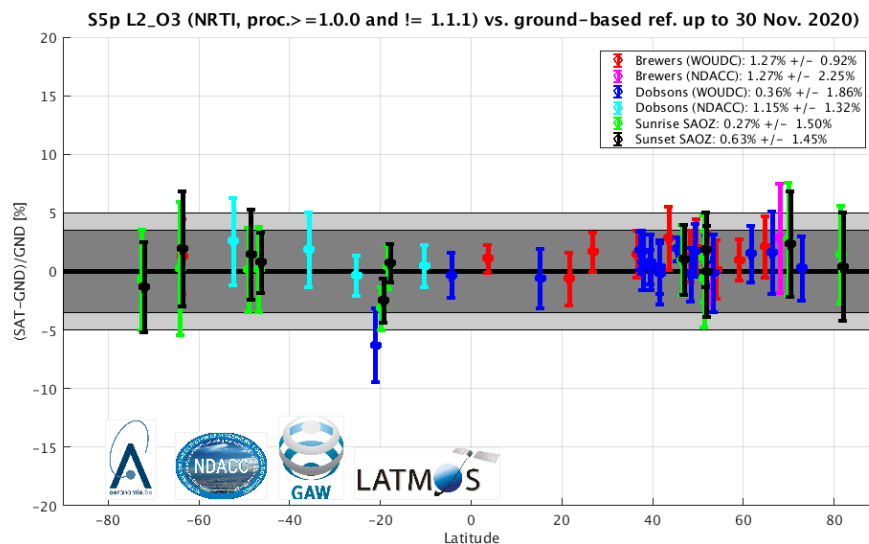


Figure 2: Meridian dependence of the median (the circular markers) and spread (± 1 sigma, the error bars) of the percent relative difference between S5p TROPOMI L2_O3 (PDGS NRTI processor v1.0.0 up to v2.1.3, 30 November 2020) and ground-based (GND) ozone column data, represented at individual stations 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).

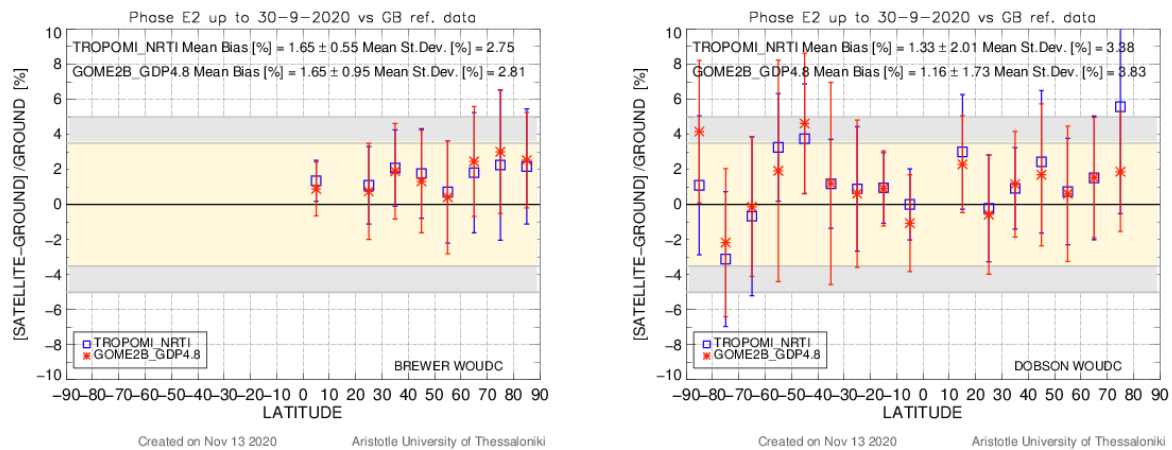


Figure 3: Comparison of the mean percentage differences between two satellite products (S5P TROPOMI L2_O3 NRTI and GOME-2B GDP 4.8) and ground-based total ozone data, versus latitude. The Brewer network comparisons are shown in the left-hand panel and the Dobson network comparisons in the right-hand panel. Both ground-based datasets are downloaded from the Woudc. Time period: is May 2018 – September 2020.

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 $5.5 \times 3.5 \text{ km}^2$, 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 (at the top in **Figure 4**), a site characterized by a complex and highly variable land/snow/sea/ice albedo, suggest an improvement since UPAS v02.01.03, as expected from replacement of the albedo climatology with the GE_LER albedo.

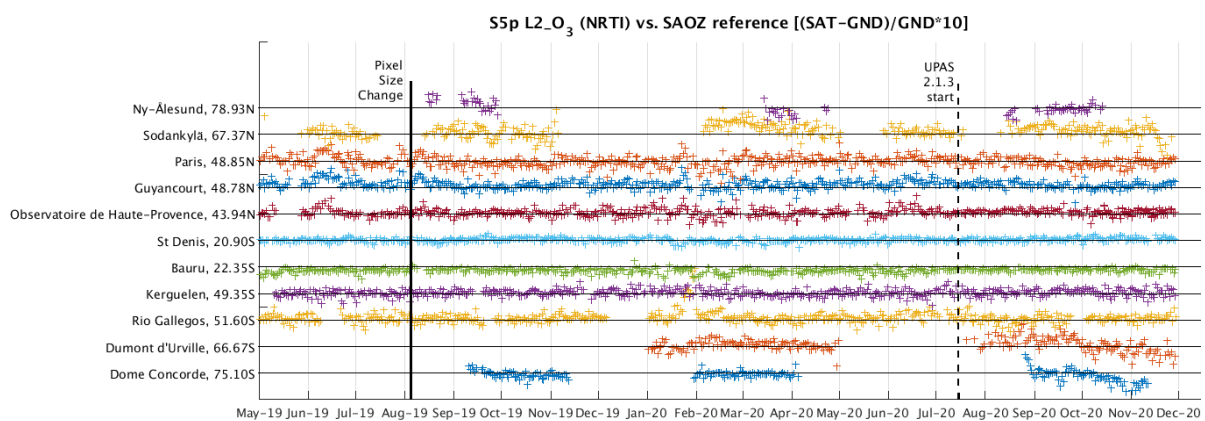


Figure 4: Time series of the difference between the S5P NRTI L2_O3 ozone column data and the SAOZ correlative data at 11 stations from pole to pole, from August 2018 to December 2020. The moment of the switch in TROPOMI pixel size (along track reduction from 7 to 5.5 km) is identified by the bold vertical line, the UPAS processor upgrade to v02.01.03 by the dashed line.

4.3.4 Dispersion

The $\pm 1\sigma$ 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 $\max.\pm 2.5\%$. 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:

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

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 1, with only a few values above 3. This suggests the ex-ante uncertainties to be realistic.

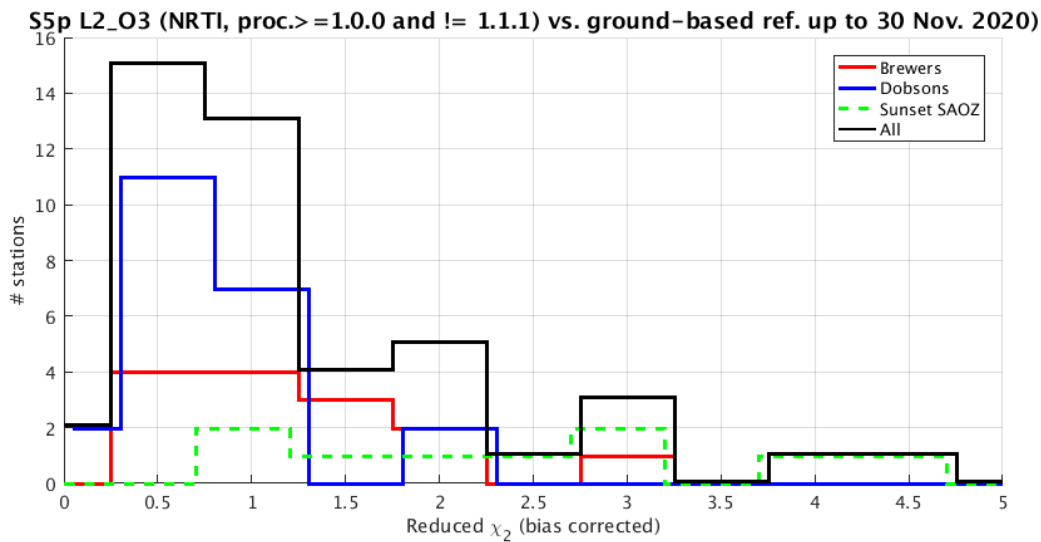


Figure 5: Histogram of the χ_r^2 values (one per reference instrument) for S5p NRTI total ozone versus different ground-based instruments.

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.

S5p L2_O3 (NRTI, proc.>=1.0.0 and != 1.1.1) vs. Brewer (WOUDC) up to 30 Nov. 2020

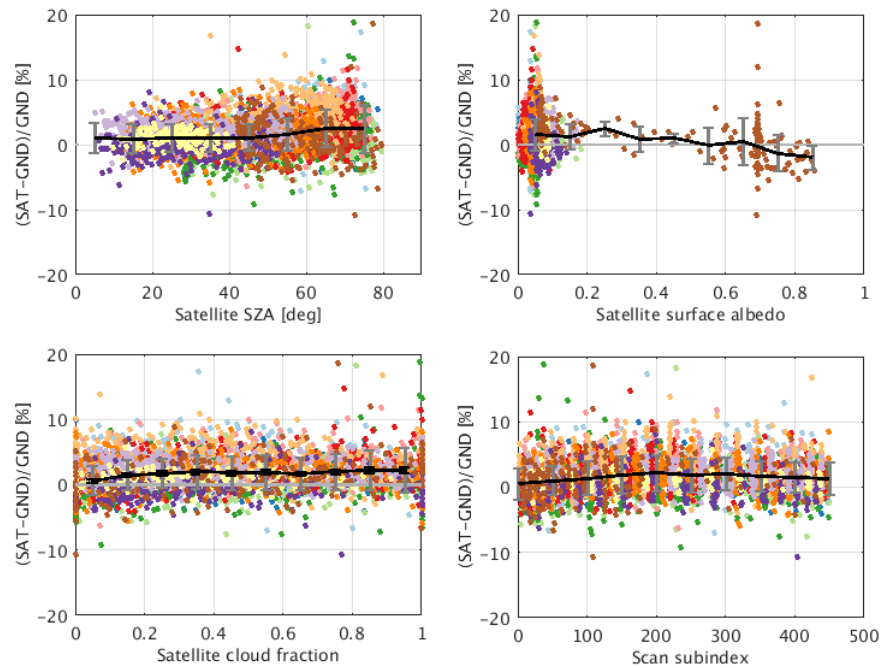


Figure 6: Dependence of the difference between S5P NRTI and ground-based Brewer total ozone data on the solar zenith angle (SZA), the surface albedo, the fractional cloud cover, and the scan sub-index of the satellite measurement. Black curve: mean and standard deviation over bins of 10 degrees of SZA,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

O3 total column at Manchester, United Kingdom

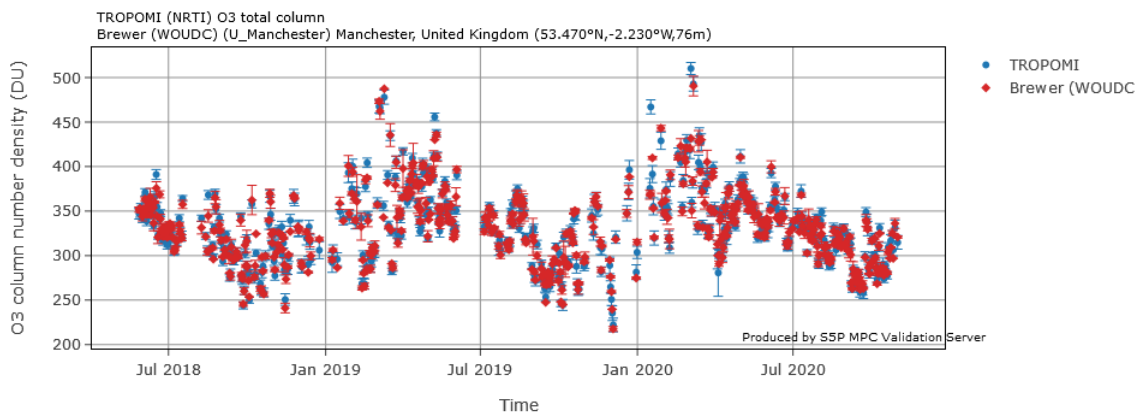


Figure 7: Time series of S5P TROPOMI NRTI and Brewer total ozone data at the station of Manchester in the United Kingdom (data courtesy J. Rimmer, University of Manchester).

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.

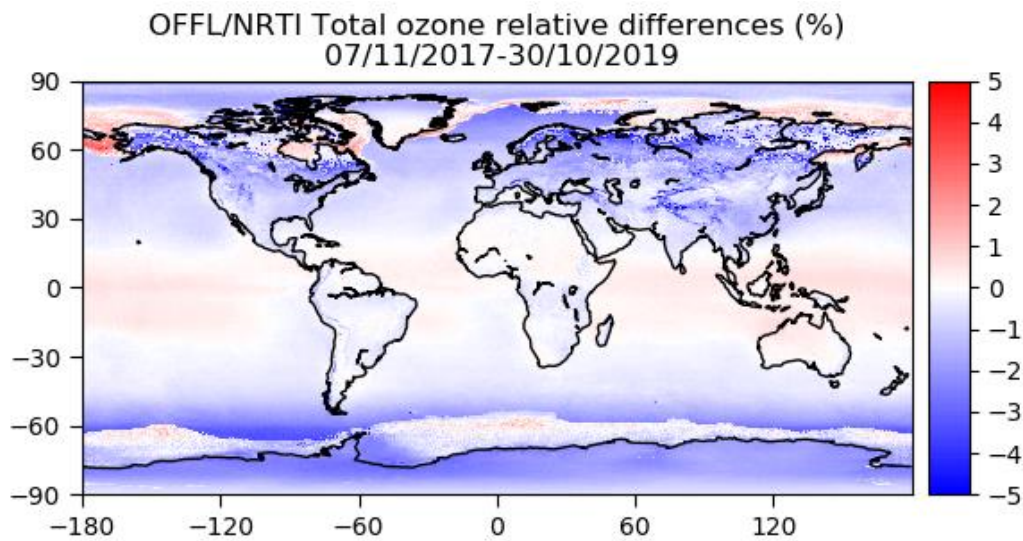


Figure 8: Percent relative difference between S5PL2_O3 total ozone data retrieved with the NRTI and OFFL processors (period November 2017 through October 2019). Geographical patterns in this case cannot be associated with real ozone features but reveal rather the effect of using either a surface albedo 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 280];
- `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>, in <https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5-precursor-workshop-2019/sentinel-5p>, and in and in <http://s5pvt.skytek.com/group/s5pvt-qwg/group-documents> 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 November 2020, and on the reference data available at the time of this report: typically, until October 2020 for the Dobson and Brewer data, and up to the end of November 2020 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**).

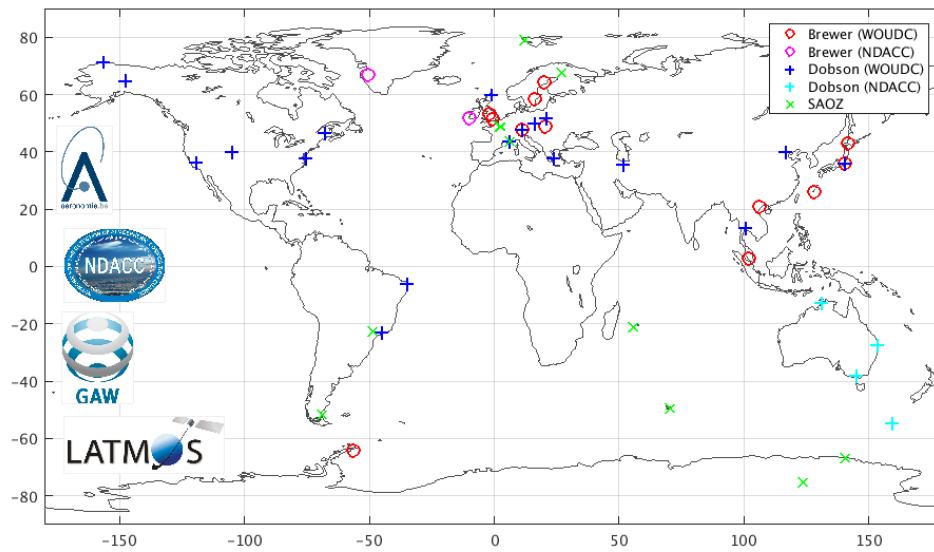


Figure 9: Geographical distribution of Brewer, Dobson and ZSL-DOAS ground-based stations for which suitable co-locations with S5P L2_O3 OFFL ozone data have been identified (period up to November 2020).

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 3.5-5%).

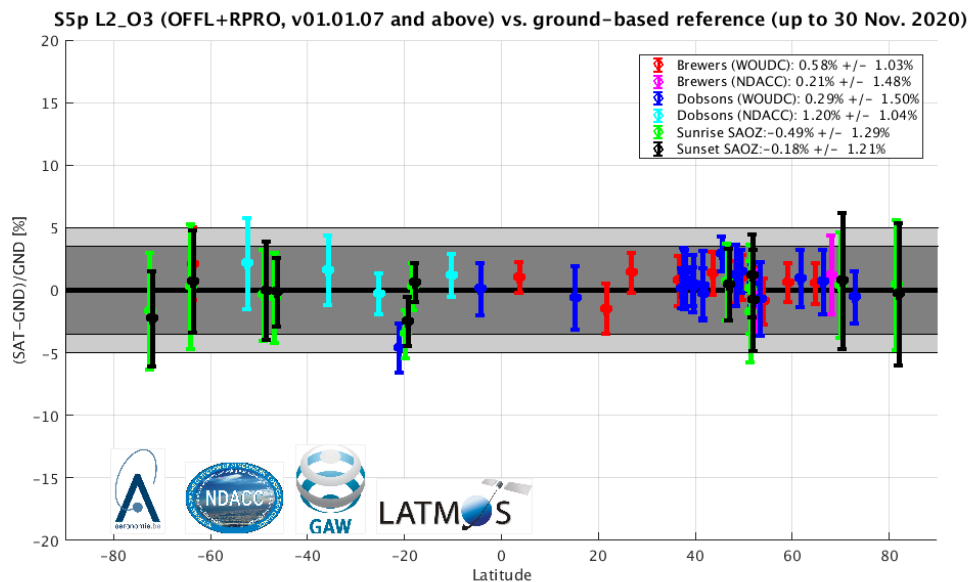


Figure 10: Meridian dependence of the median (the circular markers) and spread (± 1 sigma, the error bars) of the percent relative difference between S5P TROPOMI L2_O3 (PDGS OFFL processors v1.1.7, v1.1.8, and v2.1.3) and ground-based (GND) ozone column data, represented at individual stations from the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

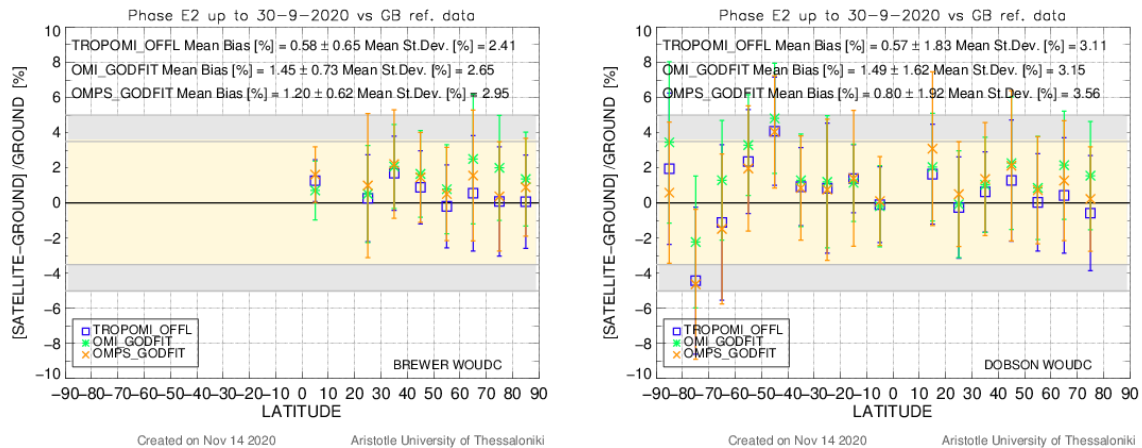


Figure 11: Comparison of the mean percentage differences between three satellite data products (S5P L2_O3 OFFL, OMI GODFIT v4 and OMPS GODFIT v4) and ground-based total ozone data, versus latitude. The Brewer network comparisons are shown in the left-hand panel and the Dobson network comparisons in the right-hand panel. Both datasets are downloaded from the Woudc. The time period of data used for these plots is May 2018 – September 2020.

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 $5.5 \times 3.5 \text{ km}^2$, 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.

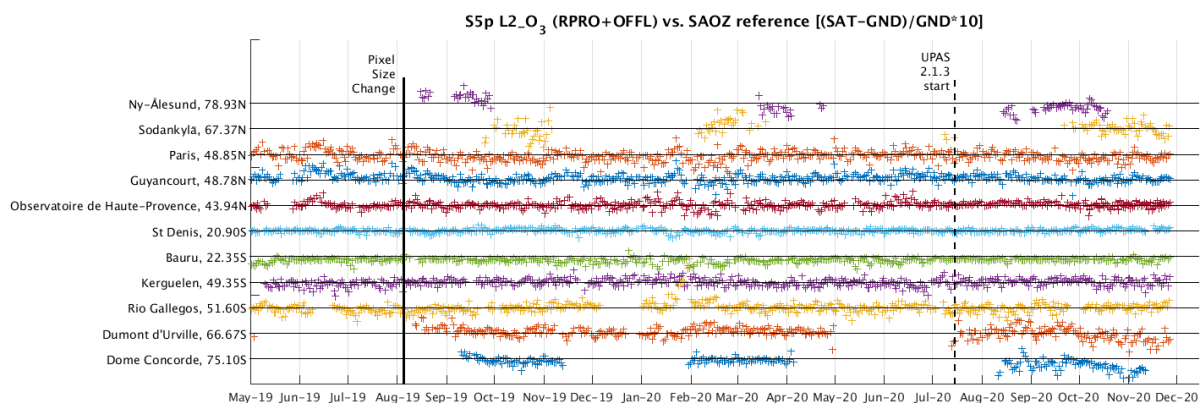


Figure 12: Time series of the difference between the S5P OFFL L2_O3 ozone column data and the SAOZ correlative data at 11 stations from pole to pole, from August 2018 to August 2020. The moment of the switch in TROPOMI pixel size (along track reduction from 7 to 5.5 km) is identified by the bold vertical line, the UPAS processor upgrade to v02.01.03 by the dashed line.

4.4.4 Dispersion

The $\pm 1\sigma$ 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:

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

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 1, with only a few values above 3. This suggests the ex-ante uncertainties to be realistic.

S5p L2_O3 (OFFL+RPRO, v01.01.07 and above) vs. ground-based reference (up to 30 Nov. 2020)

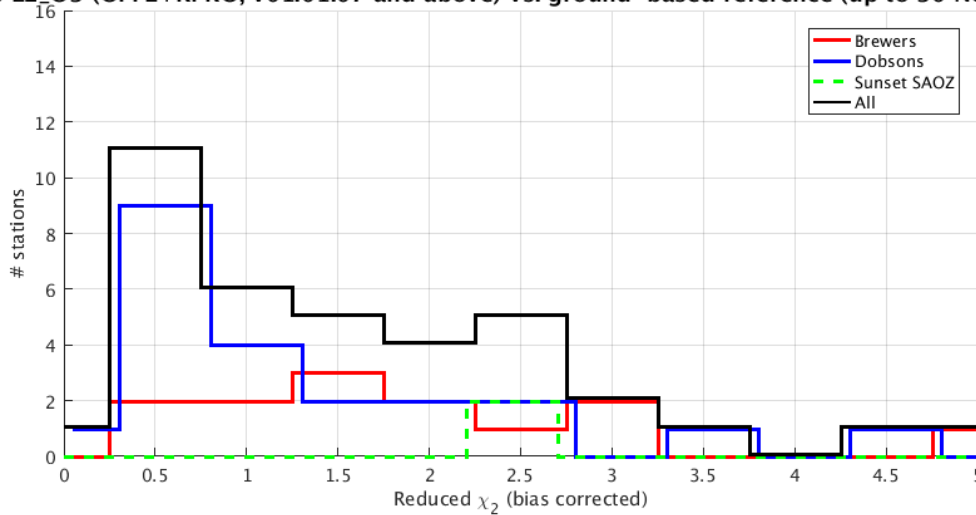


Figure 13: Histogram of the χ_r^2 values (one per reference instrument) for S5p NRTI total ozone versus different ground-based instruments.

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 16**. Also, there is a modest increase in scatter for measurements at larger cloud fractions.

S5p L2_O3 (OFFL+RPRO, v01.01.07 and above) vs. Brewer (WOUDC) up to 30 Nov. 2020

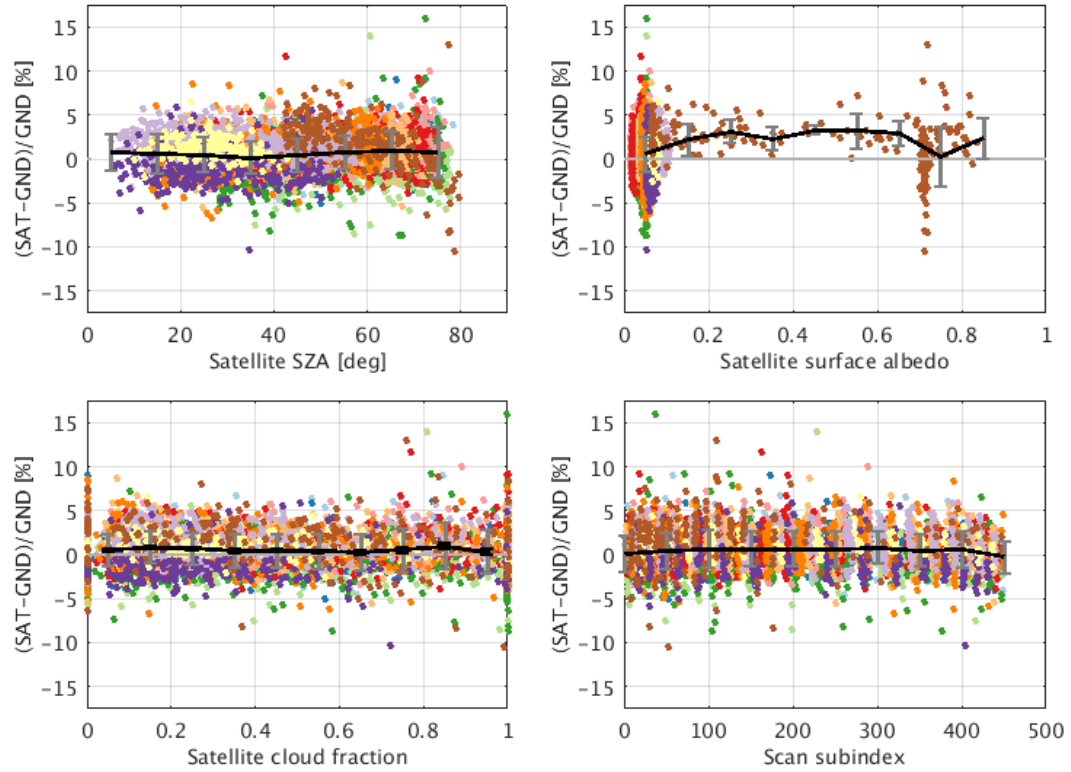


Figure 14: Dependence of the difference between S5P OFFL and ground-based Brewer total ozone data on the solar zenith angle (SZA), the surface albedo, the fractional cloud cover, and the scan sub-index of the satellite measurement. Black curve: mean and standard deviation over bins of 10 degrees of SZA,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

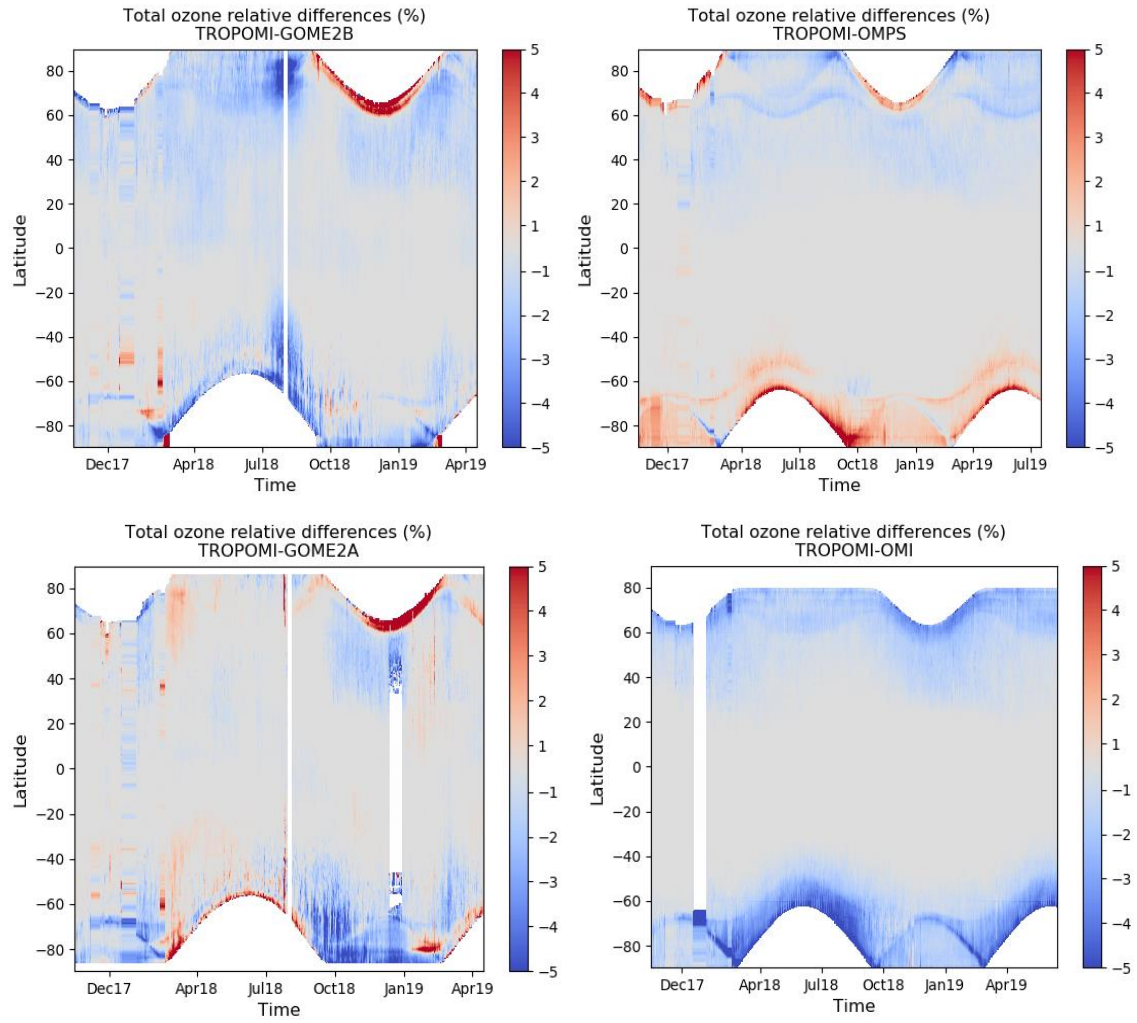


Figure 15: Latitude and time cross-section of the percent relative difference between S5P L2_O3 OFFL total ozone data and comparable total ozone data sets from other satellite: GOME-2B GDP 4.8, S-NPP OMPS, GOME-2A GDP 4.8, and Aura OMI.

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.

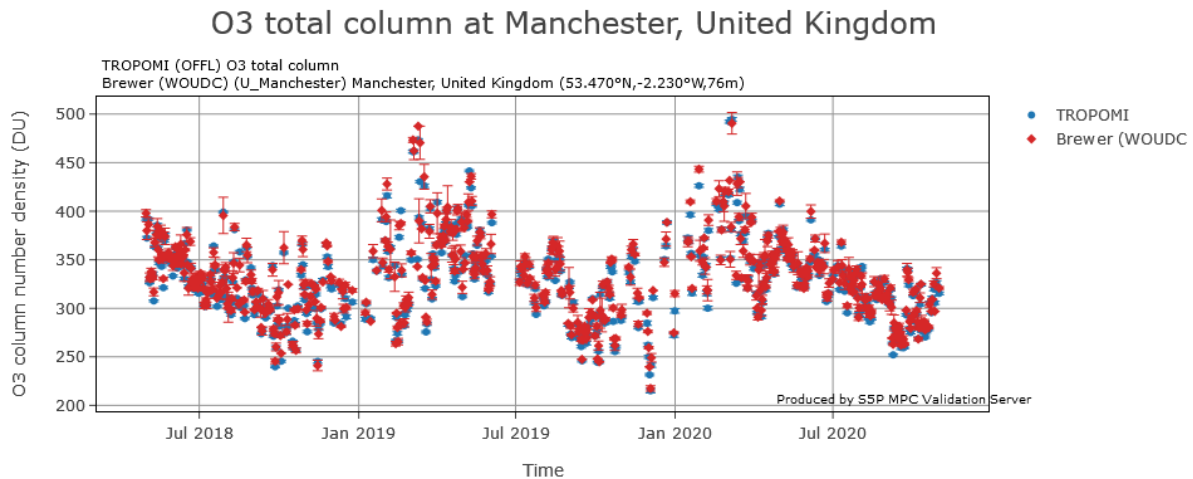


Figure 16: Time series of S5P TROPOMI OFFL and Brewer total ozone data at the station of Manchester in the United Kingdom (data courtesy J. Rimmer, University of Manchester).

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 – 10 November 2020, the ground-based validation analysis considers 909 S5P OFFL CCD data products and 632 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 – July 2020) and to OMI data (May 2018 – September 2020).

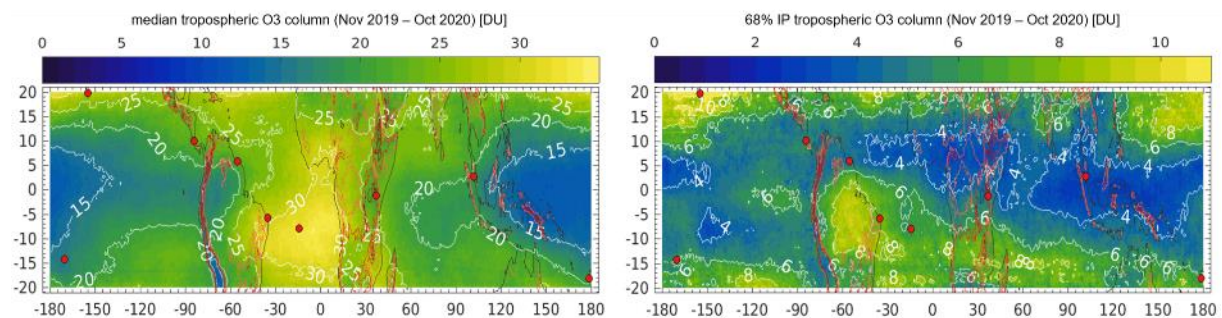


Figure 17: Median value (left) and half-width of 68% interpercentile (right) of S5P OFFL tropospheric ozone column data (CCD) over a full year of operations (November 2019 – October 2020). Red markers locate the nine ground-based ozonesonde stations used in the validation analysis. Red contours indicate surface elevation (500, 1000, 2000 m). These maps provide context to **Figure 18** and **Figure 19**.

5.3.3 Bias

S5P tropospheric O₃ 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 +15% 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 +0.1 DU or +0.3%. The larger mean difference of +2.4 DU or +14% 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 16-32% or 3.7-6.9 DU (**Figure 19** and **Figure 21**), and the network average is 24% or 4.6 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 O₃ field and there is a considerable difference in spatio-temporal sampling between S5P and ozonesonde. In addition, the random component of the uncertainty of the ozonesonde measurement contributes about 5-10% to the observed dispersion in the differences. Hence, the uncertainty of the S5P data is better than the 24% 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, which is coherent with the average dispersion found in comparisons to the ground-based network. 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 seems to appear in the difference between S5P and both other satellite data records (**Figure 20**). The sampling of the ground-based comparisons is too limited at this point to confirm a seasonal cycle. The oscillation is clear in 2019 and persists until April 2020. The phase appears to vary with latitude; 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). However, in June-July 2020 the difference time series show a local minimum where a maximum was expected. This change in behavior already starts in April-May and can therefore not be related to the switch in processor version early July. This will be monitored in the coming months.

S5P time series in **Figure 18** show signatures of biomass burning affected air (high tropospheric O₃ values) over Atlantic and African sites (Heredia, Paramaribo, Natal, Ascension Island, Nairobi), and of intense convective activity (low tropospheric O₃ values) over Pacific stations (Samoa, Suva, Sepang Airport). During the 2018 biomass burning season the positive S5P bias w.r.t. Paramaribo, Heredia and Nairobi is clearly larger than during the rest of the year. The temporary, additional bias amounts to, on an average, about 25% or 5 DU. A period of larger positive biases reoccurred at Paramaribo in the 2019 and 2020 seasons and to a lesser extent at Natal and Ascension Island as well. However, the length of the elevated bias period for the latter two sites does not cover the entire peak of high tropospheric O₃ conditions. The relation between an elevated, positive bias for S5P and biomass burning is therefore not clear and subject of further investigation. 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 73% at individual stations, while the network average is 60% (**Figure 21**, top left).

5.3.7 Geographical patterns

Annual median TROPOMI data (November 2019 – October 2020, **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 orbital sampling by the S5P instrument progression imprints another random error pattern (up to 5 DU) that follows the progression of S5P orbit. These errors are correlated in time and space and appear at small spatio-temporal scales.

Other known geophysical patterns and oscillations, such as the annual and semi-annual cycles, the biomass burning season and the Madden-Julian Oscillation, are present in the S5P tropospheric O₃ data record as well. 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 are generated using the new processor. By the time of this report 87 products were processed using 02.01.03, leading to 58 co-locations with ozonesondes at four sites (Hilo, Paramaribo, Samoa, Suva). No clear change was seen in the comparisons (**Figure 19**) although confidence is currently limited due to the small number of comparisons in the latter period in combination with the considerable dispersion in the comparisons. More precise assessments will be possible with longer time series in the next update of this report.

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.

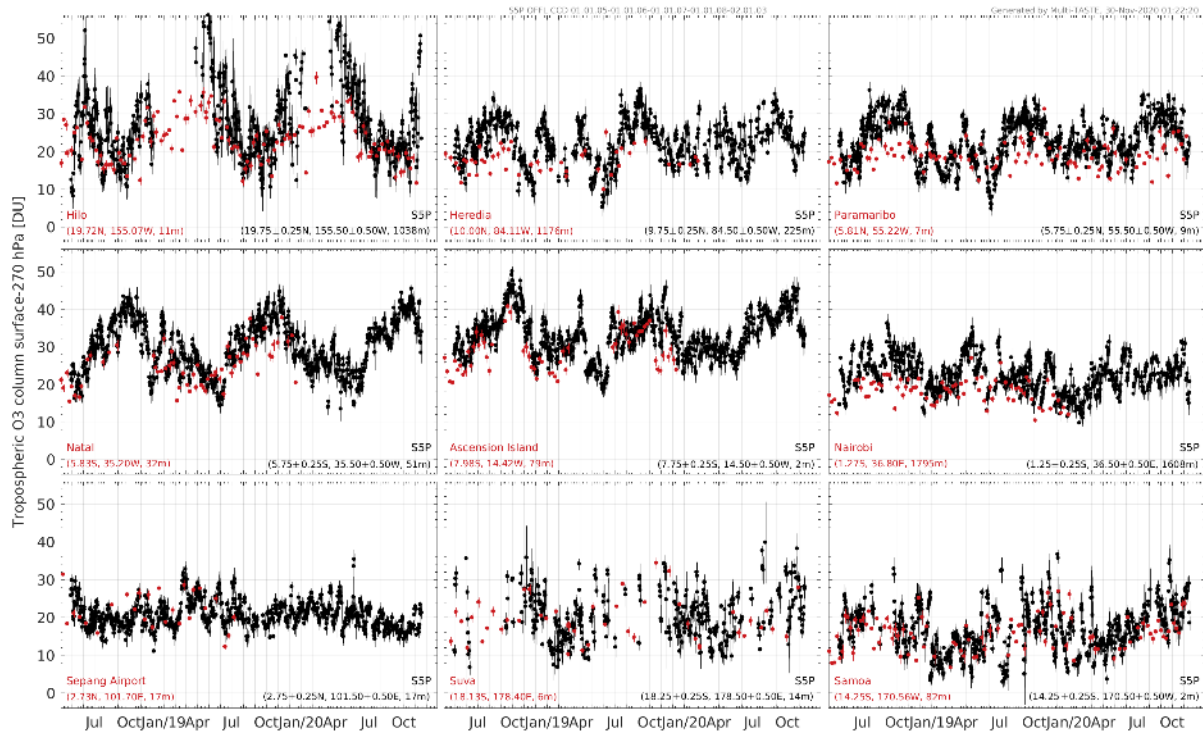


Figure 18: Time series of spatially co-located tropospheric O₃ column data by ozonesonde (red) and by S5P OFFL v01.01.05-v02.01.03 (black). All data were screened following recommendations by the data providers.

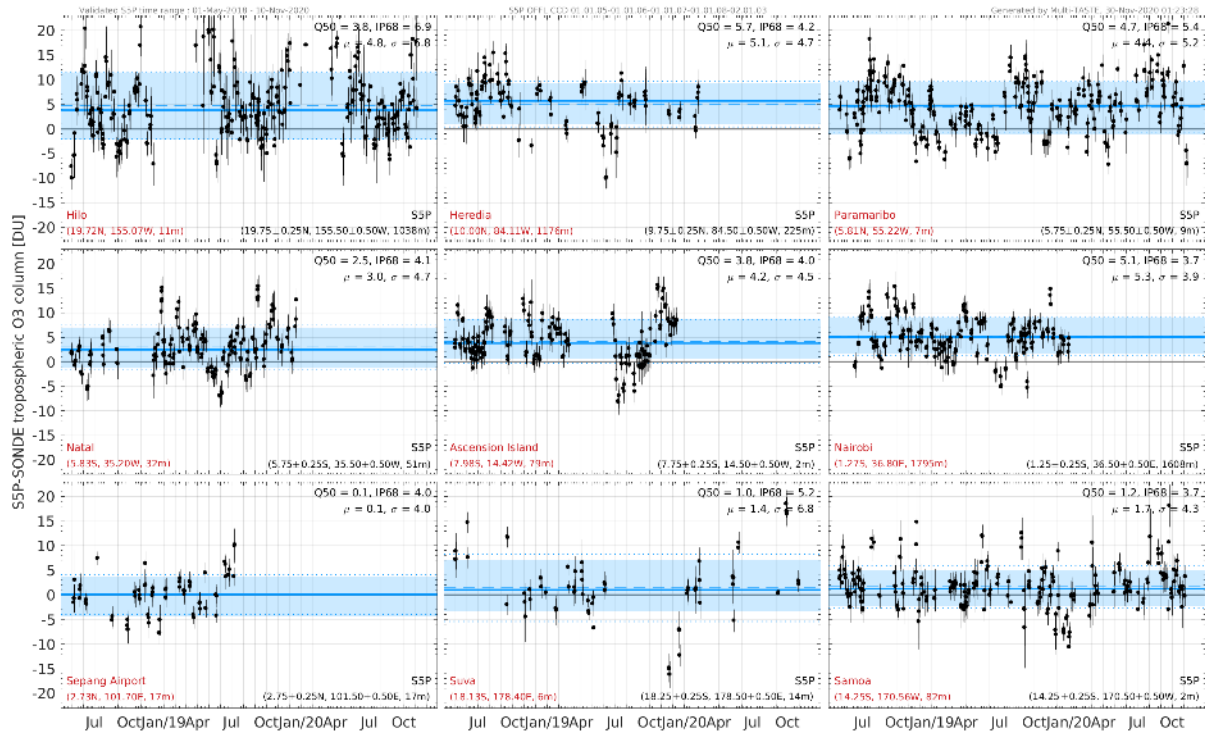


Figure 19: Time series of the absolute difference between spatially and temporally co-located S5P and ozonesonde tropospheric O₃ column data. The blue line and shaded area shows the median value and the range between the 16% and 84% percentiles. Positive values indicate a high bias of S5P w.r.t. the reference.

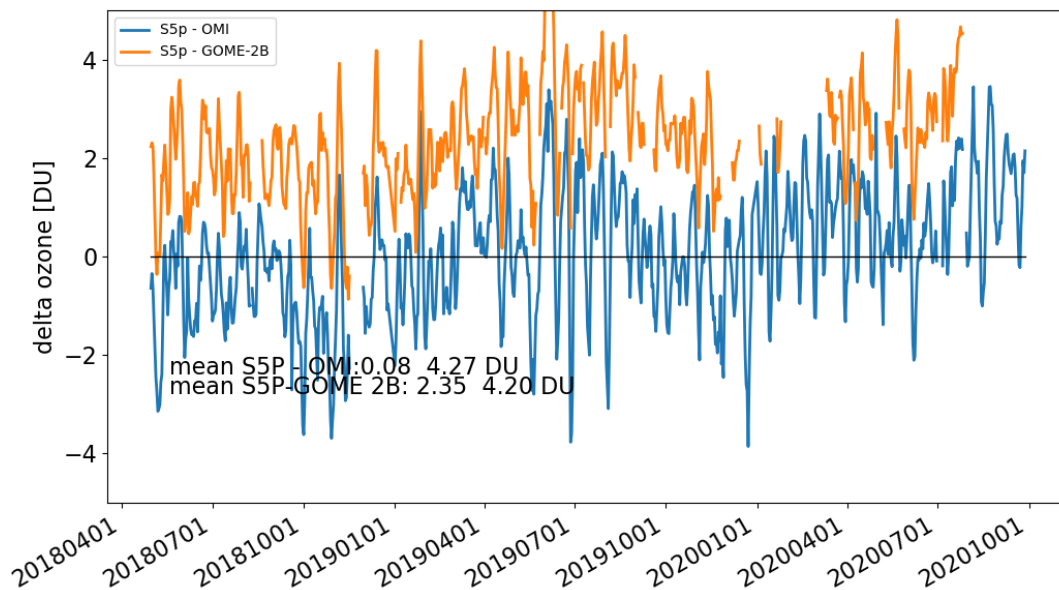


Figure 20: Difference time series of daily tropospheric O₃ column data averaged over the 20°S – 20°N tropical belt. S5P OFFL CCD data are compared to satellite data by OMI and GOME-2B, positive values indicate a high bias of S5P w.r.t. the reference.

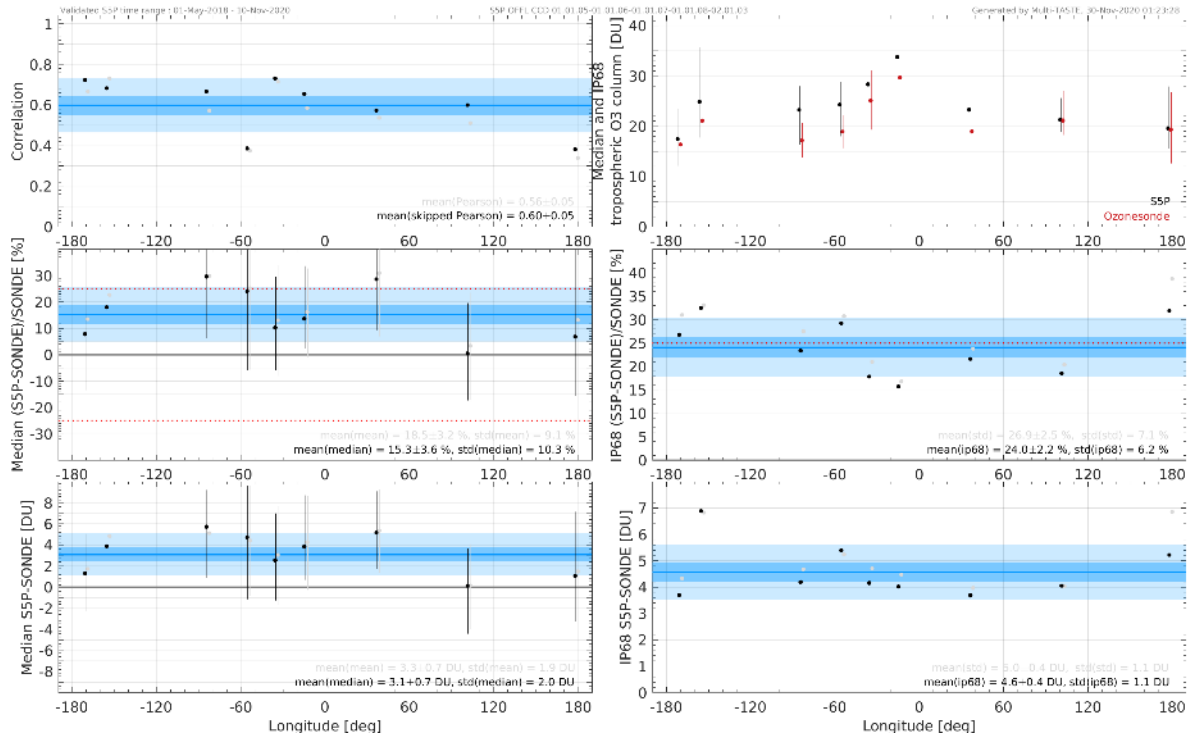


Figure 21: Overview of correlation (top left), median bias (middle & bottom left) and dispersion of the difference (middle & bottom right) of S5P tropospheric O₃ column data for each SHADOZ site (black markers). Black vertical bars represent the 68% interpercentile of the difference. The mean, standard 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 NO₂ stratospheric column, the NO₂ tropospheric column, and the NO₂ total column. Validation results are discussed with respect to the product quality targets outlined in **Table 3**. The OFFL product was reprocessed to version 01.02.02. The current processor version is 01.03.02 for NRTI and OFFL.

OFFL and NRTI products may differ because are retrieved using ECWMF forecast meteorological data with a time difference as input for the CTM. Subsection 0 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

Stratospheric NO₂ – 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 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.

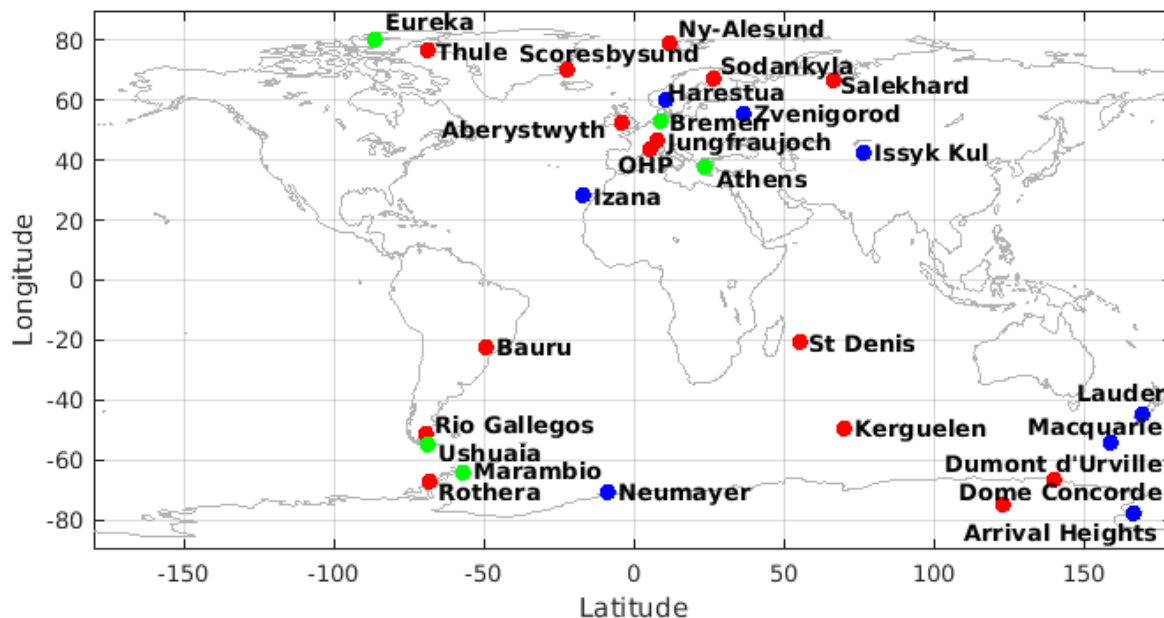


Figure 22: Geographical distribution of the NDACC ZSL-DOAS instruments measuring routinely stratospheric NO₂ and yielding co-locations with the current S5P L2_NO2 datasets. Stations marked with a red dot contribute fast delivery data coming from the LATMOS_RT facility. Blue and green dots depict the 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 NO₂ 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 NO₂, the ZSL-DOAS measurements 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 taken here 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 NO₂ 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 also 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 Alzomoni (3985 m), Izaña (2360 m), and Mauna Loa (4169 m).

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 data are collected through the S5PVT AO Project NIDFORVAL, while a few PIs also contribute data from seven MAX-DOAS stations to the ESA Atmospheric Validation Data Centre (EVDC). The VDAF Automated Validation Server (VDAF-AVS) performs comparisons for the stations of Athens and Bremen (IUP-B), Cabauw and De Bilt (KNMI), Mainz (MPIC), Uccle and Xianghe (BIRA-IASB).

MAX-DOAS tropospheric NO₂ 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., AMT, 2020). 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 the Xianghe site to the VDAF-AVS and several time-period updates.

Careful revision (and eventual reprocessing) are ongoing for the Thessaloniki site and the UNAM sites (Cuautitlan, Unam and Vallejo), while complete years are submitted by the JAMSTEC team (Cape Hedo, Fukue, Yokosuka). Contacts with new stations PIs (Mohali, Munich, Melbourne...) have led to the inclusion of the Mohali indian site in this report.

Total NO₂ – Pandora Direct-Sun UV-Visible Spectrometers

S5P 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 data have a `qa_value` > 0.5;
- the TROPOMI ground pixel contains the Pandora station;
- Pandora measurements with a flag not equal to 0 or 10 and negative values are excluded;
- the closest Pandora measurement 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 NO₂ can also be representative of the stratospheric NO₂ column.

6.2.2 Satellites

S5P 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, 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 at the current stage to only use TROPOMI NO₂ pixel associated with a `qa_value` above 0.75. This removes cloudy scenes (cloud radiance fraction > 0.5), scenes covered by snow/ice, several other errors, and problematic retrievals. For stratospheric NO₂ retrievals and data comparisons, clouds are less of a problem, so that `qa_value` above 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).

The NRTI data covers the full range of versions from 01.00.01 to 01.03.02 as there is no reprocessing necessary. The OFFL data has been reprocessed to version 01.02.02 from 2018-05-01 until 2018-10-17. After that, the processor version has been changed five times from 01.02.00 up to 01.03.02, which is the current version, started in 2019-06-26.

This summary takes into consideration (updates of) the results reported at the *S5P First Public Release Validation Workshop* (ESA/ESRIN, June 25-26, 2018) and at the *3rd S5PVT workshop* (ESA/ESRIN, November 11-14, 2019). The contributions are archived at nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5p-first-product-release-workshop/sentinel-5p and nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5-precursor-workshop-2019/sentinel-5p. Up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal at mpc-vdaf.tropomi.eu. An in-depth discussion of the routine validation results up to March 2020 is now published (Verhoelst *et al.*, 2020).

PGN steadily extends its Pandora data provision on EVDC, making it accessible for the VDAF Automated Validation Server. Furthermore, there is ongoing 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. The FRM data streams that are stored at the VDAF-AVS are:

- NDACC ZSL-DOAS stratospheric columns from 25 stations are used, sampling the latitude range from 80°N (Eureka) to -75°S (Dome C) and covering the period from May 2018 to November 2020.
- PGN total column direct-sun data available at the currently cover 26 sites. Co-locations cover the time period May 2018 to November 2020.
- MAX-DOAS from the NIDFORVAL AO project from 21 stations (new: Mohali) were used, covering the time period 2018-04-30 to 2020-11-07. Vertical smoothing could be applied for 12 stations.
- NDACC MAX-DOAS tropospheric column data are available from 7 stations (new: Xianghe) are available at the VDAF Automated Validation Server for the time period from May 2018 until November 2020.

6.3.3 Stratospheric NO₂ column

6.3.3.1 Bias

S5P NRTI stratospheric NO₂ column values are generally lower than ground-based values by approximately -0.17 Pmolec/cm², with a station-to-station scatter of the mean bias of similar magnitude (**Figure 23**). The validation results are based on comparisons at 25 NDACC ZSL-DOAS stations, sampling the latitude range from 80°N (Eureka) to -75°S (Dome C). The median bias for the OFFL stratospheric column is -5.5% (-0.2 Pmolec/cm²) with a station-to-station scatter of 11%, using data from 13 VDAF-AVS ZSL-DOAS stations (7129 co-locations). Thus both datasets are within the mission requirement of 10% maximum bias (equivalent to 0.2-0.4 Pmolec/cm², depending on latitude and season). The ZSL-DOAS stratospheric NO₂ column values measured at twilight are converted to the measurement time of TROPOMI using the aforementioned PSCBOX 1D stacked-box photochemical model initiated by SLIMCAT fields.

6.3.3.2 Dispersion

The $\pm 1\sigma$ dispersion of the difference between stratospheric column and reference data around their mean value rarely exceeds 0.3 Pmolec/cm² at sites without tropospheric pollution (cf. the error bars in **Figure 23**). 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.

6.3.3.3 Dependence on influence quantities

Potential dependence of the S5P stratospheric column bias and dispersion on the Solar Zenith Angle (SZA), cloud fraction (CF) and surface albedo of the S5P measurement is evaluated. The evaluation does not reveal any variation of the bias much larger than 0.4 Pmolec/cm² over the range of the influence quantities (**Figure 24**).

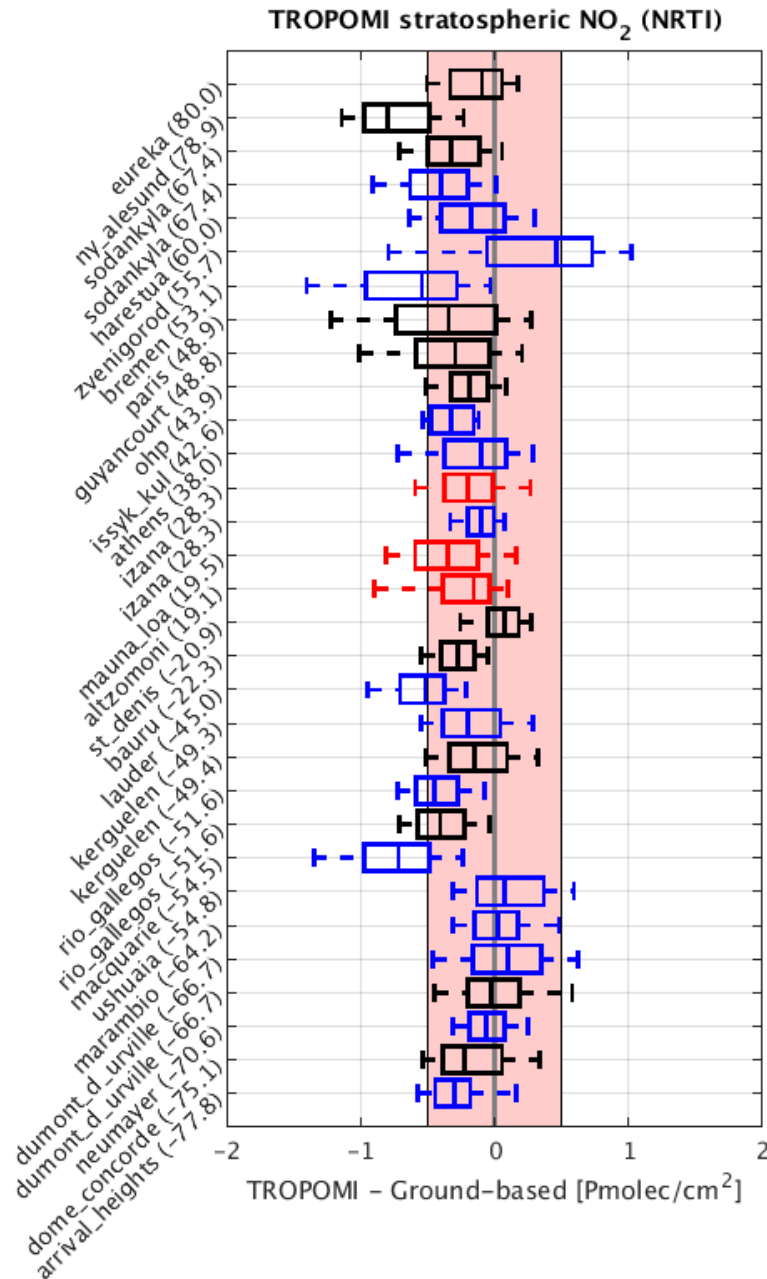


Figure 23: Box-and-whisker plots summarizing the pole to pole bias and spread [Pmolec/cm²] of S5p TROPOMI NRTI and NDACC ZSL-DOAS NO₂ stratospheric columns (SAOZ data in black, other ZSL-DOAS in blue, and mountain-top PGN in red). The median difference is represented by a vertical solid line inside the 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 of 0.5 Pmolec/cm² for the uncertainty. Station name and latitude of the station are added on the left. The network-wide mean bias and its formal uncertainty are -0.17 and 0.06 Pmolec/cm², 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 October 2020.

6.3.3.4 Seasonal cycle and shorter term variability

S5P and ground-based ZSL-DOAS instruments capture similarly the short-term variability (at daily and monthly scales) of the NO₂ stratospheric column, as illustrated at the NDACC station of Kerguelen Island in **Figure 25**. 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.

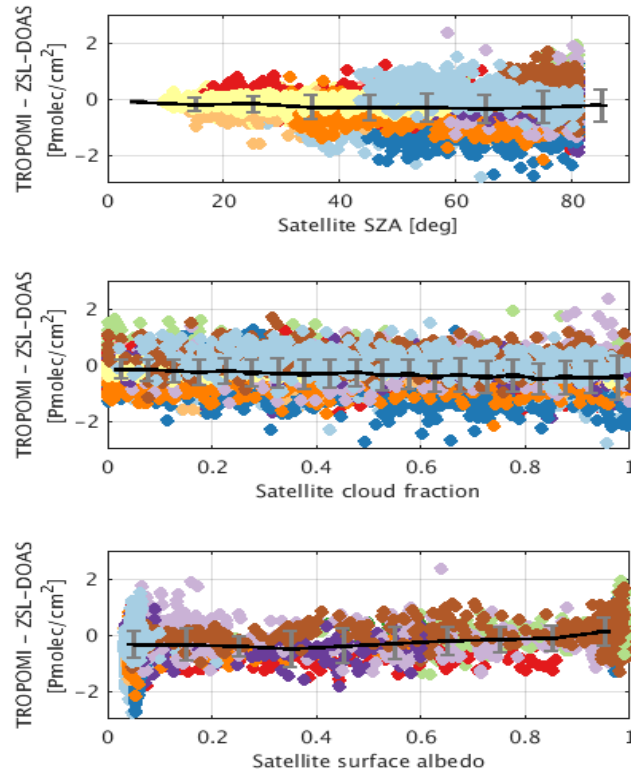


Figure 24: Difference between S5P L2_NO2 NRTI and ground-based SAOZ stratospheric NO₂ column data as a function of the satellite solar zenith angle (SZA), satellite cloud fraction and satellite surface albedo. Mean and standard deviation calculated over bin widths of 10 degrees in SZA, 0.1 in CF, and 0.1 in surface albedo.

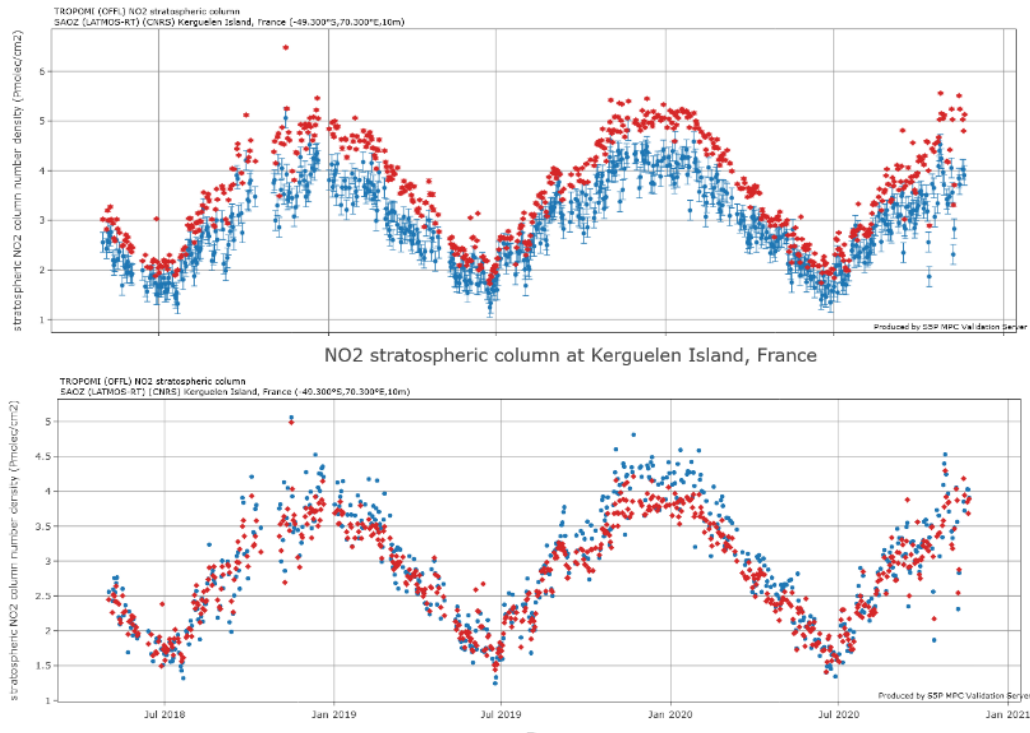


Figure 25: Time series of S5P OFFL L2_NO2 stratospheric NO₂ column data (blue dots) co-located with ground-based SAOZ twilight measurements (red dots) at sunset performed by LATMOS at the NDACC southern mid-latitude station of Kerguelen Island. In the upper plot, the photochemical correction is deactivated to offset the two time series and better see the day-to-day variability. The time frame is July 2018 to November 2020.

S5P stratospheric NO₂ and Pandora total NO₂ plotted as a function of the calendar month at 3 mountain sites follow the same seasonal cycle (**Figure 26**). It must be noted that the PGN Pandora data (even after removing negative Pandora values as done here) is more scattered. The 30-day rolling median of the relative difference is within the bias requirements, except for the months January-March at Mauna Loa, which seems to be due to elevated Pandora values.

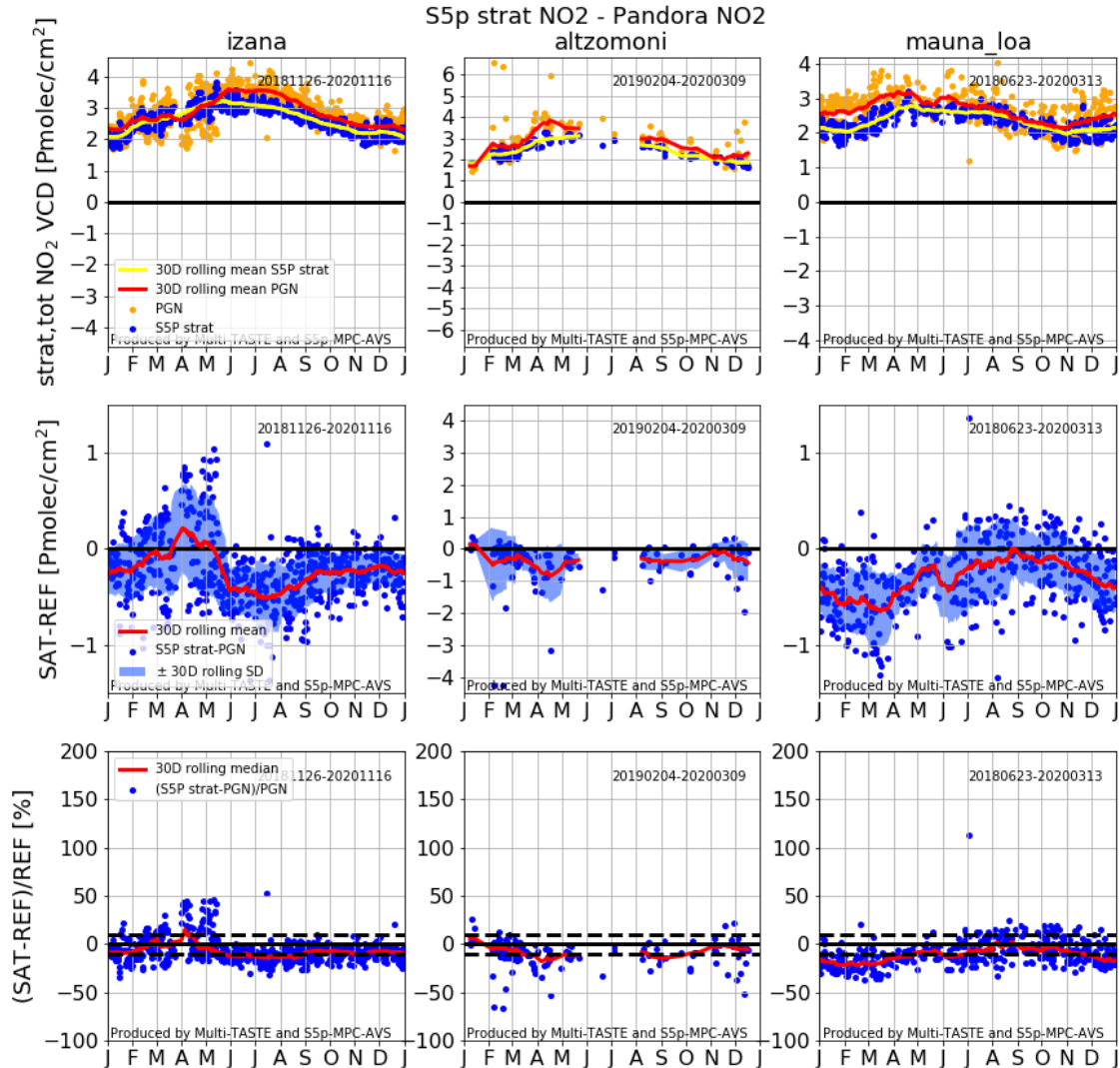


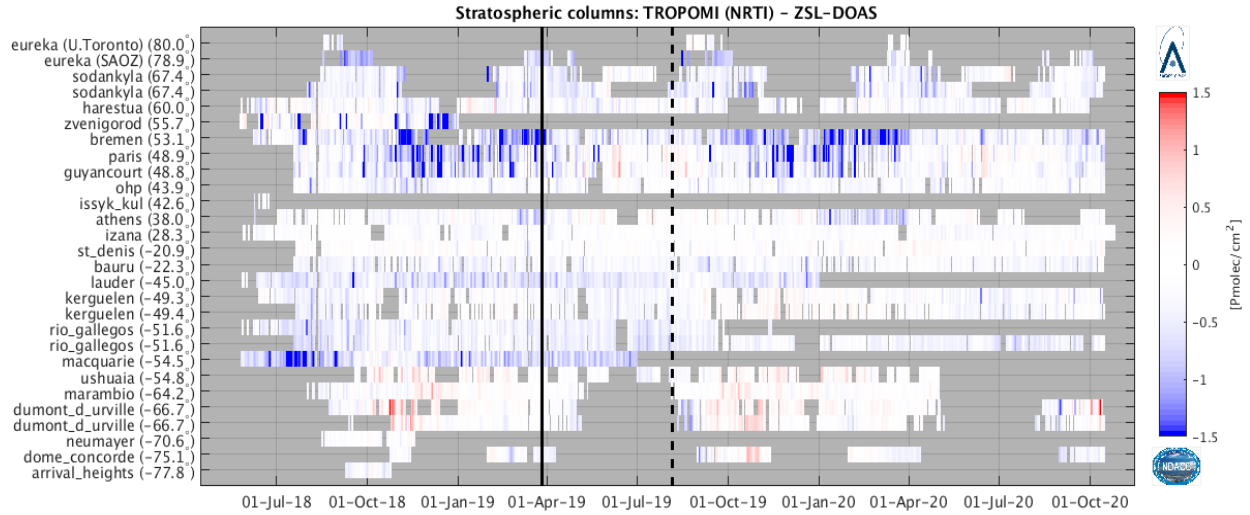
Figure 26: Seasonal cycle plots of the S5P stratospheric NO₂ column and Pandora total column at 3 mountain sites. Upper: S5P and PGN Pandora together; middle: their difference; lower: their percent relative difference. Plain lines represent the 30-day rolling mean or (for relative difference) the 30-day rolling median (cyclic boundary conditions). The data is mapped to one generic year and obtained from validation server on 2020/11/26.

6.3.3.5 Geographical patterns

None to report.

6.3.3.6 Processor version change and switch to smaller ground pixel resolution

The effect of (1) the upgrade of the NL-L2 processor to v1.3.0 on 20 July 2019 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 27**) and 3 PGN Pandora mountain sites (**Figure 28**). The difference between the S5P and ground-based data does not show any impact from processor and pixel size change.



Update of Verhoelst et al., AMTD 2020–119, 2020, produced by BIRA–IASB for the S5p MPC VDAF on 03–Nov–2020 using NDACC reference data

Figure 27: Time series – from August 2018 until October 2020 – of S5P NRTI L2_NO2 and NDACC ZSL-DOAS NO₂ stratospheric column differences [Pmole/cm²]. 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. Data processed both in NRT mode (LATMOS_RT) and with the LATMOS_v3 processor for stations appearing twice. Both are shown in order to illustrate the consistency.

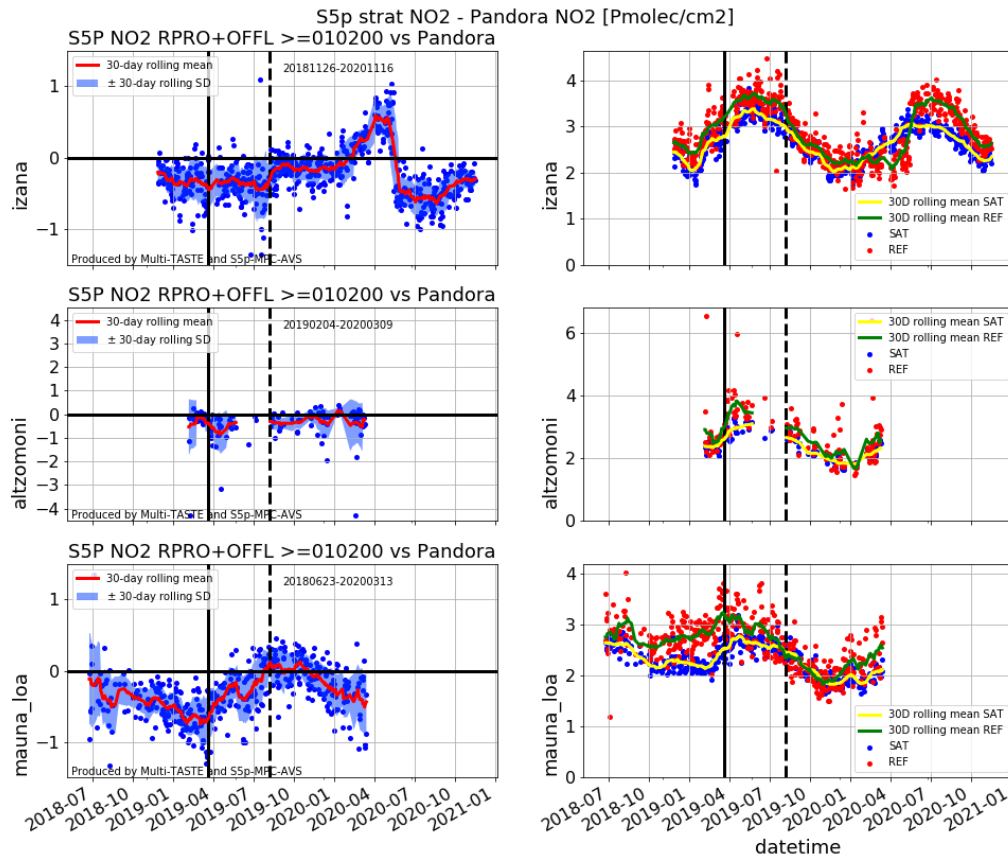


Figure 28: Time series of the difference (left) between S5P RPRO+OFFL L2_NO2 NO₂ stratospheric columns and PGN Pandora total NO₂ columns at 3 mountain sites. The S5P and Pandora time series are on the right. Full vertical line: OFFL processor version change on 2019/03/20. Dashed line: pixel size switch at 2019/08/06. Data covers the period June 2018 until November 2020 and is obtained from the validation server on 2020/11/26.

6.3.3.7 Other features

None to report.

6.3.4 Tropospheric NO₂ column

To contain the global outbreak of the COVID-19 pandemic, lockdown measures affecting transport and production were taken that had an, with impact on tropospheric NO₂ concentration levels. This was also observed world-wide by TROPOMI (e.g. Bauwens *et al.*, 2020). **Figure 29** presents the tropospheric NO₂ column density at Athens as observed by S5P TROPOMI and by the MAX-DOAS instrument operated by IUP-B. First nation-wide measures (closure of schools and universities) started on March 10, 2020, with further restrictions on later days (closure of business, ban on non-essential movement, see <https://www.bloomberg.com/news/articles/2020-04-17/humbled-greeks-show-the-world-how-to-handle-the-virus-outbreak>). Both MAX-DOAS and S5P observe a significant drop NO₂ between March 3 and March 13 (no co-locations in between these two dates). NO₂ column values stay consistently low for weeks after. It can also be observed that low NO₂ column values are better reproduced (in absolute scale) than high NO₂ column values, where S5P tends to underestimate compared to MAX-DOAS observations. This is consistent with the validation results reported in Section 6.3.5.1.

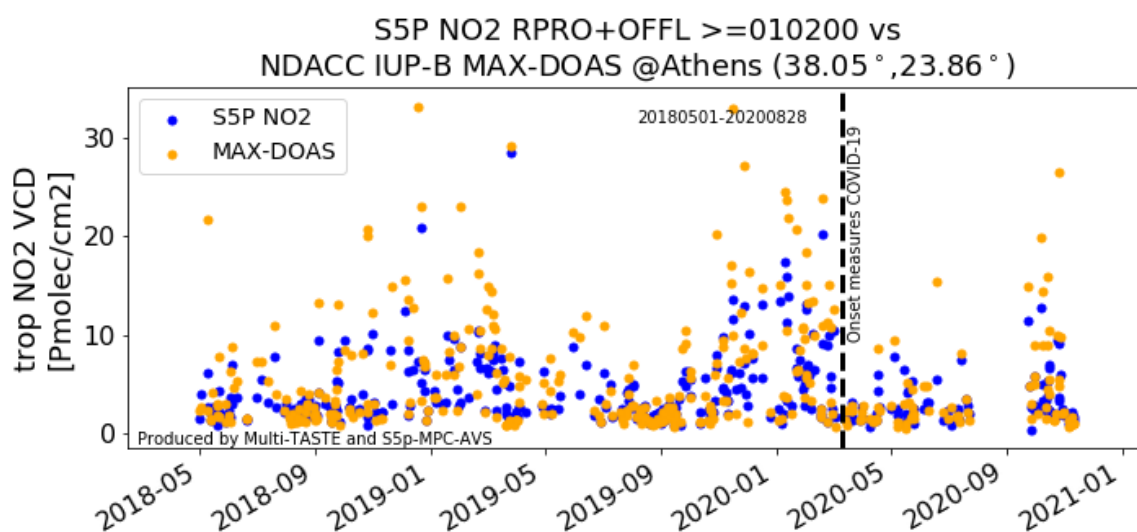


Figure 29: Tropospheric NO₂ column at Athens, as observed by S5P TROPOMI and MAX-DOAS. First nation-wide measures (closures of schools and universities) to contain COVID-19 started on March 10 2020 (dashed line), with further restrictions on later days (source: <https://www.bloomberg.com/news/articles/2020-04-17/humbled-greeks-show-the-world-how-to-handle-the-virus-outbreak>). Note how both TROPOMI and MAX-DOAS capture the drop in tropospheric NO₂ and return to normal values in summer.

6.3.4.1 Bias

S5P OFFL NO₂ tropospheric column values are compared with data from 21 NIDFORVAL MAX-DOAS stations. The median bias over all stations is -2 Pmolec/cm² (-36.5%). **Figure 28** and **29** show time series and biases for these stations. Using only data from the 7 VDAF-AVS MAX-DOAS stations mainly from Europe (latest inspection on 2020-11-25, 2598 co-locations), the median bias reduces to approximately -1.3 Pmolec/cm² (-25%). The bias varies between -0.2 Pmolec/cm² and -2.4 Pmolec/cm² with a median relative difference between -8% (Athens) and -36% (Cabauw, Bremen, Mainz). These results are within the mission requirement of a maximum bias of 50%.

As the site-to-site dispersion (IP68/2 over all site medians) is 3.3 Pmolec/cm², a single bias number for all sites has limited meaning. As discussed in Verhoelst (2020), three regimes are identified in the comparisons: (1) small tropospheric NO₂ 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 the 3 regimes is about 0.5 Pmolec/cm² (7%), -2.2 Pmolec/cm² (-36.8%) and -13 Pmolec/cm² (-54%), respectively.

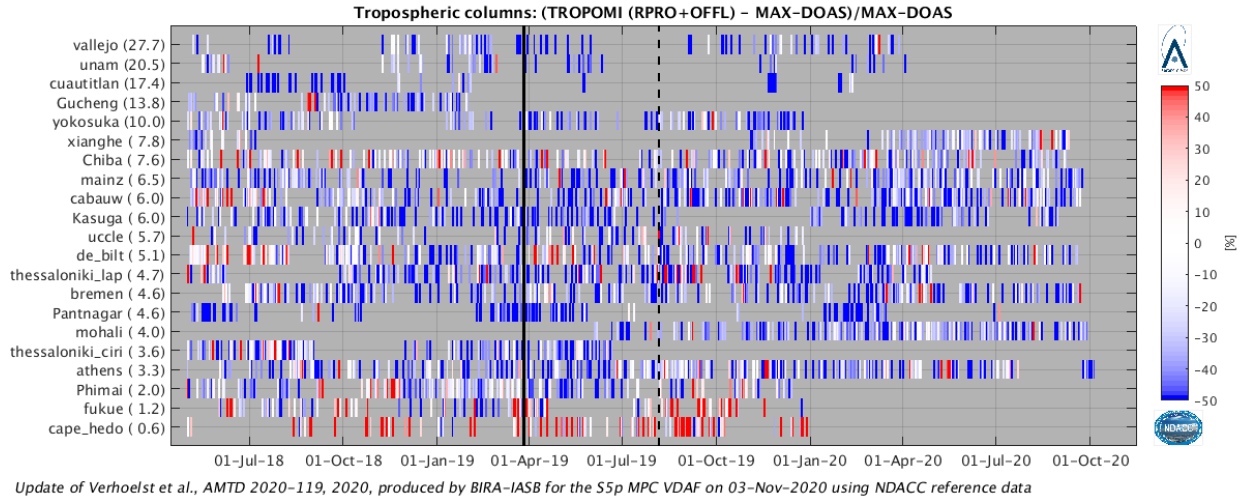


Figure 30: Time series – from May 2018 until October 2020 – of the difference between S5P RPRO+OFFL and MAX-DOAS NO₂ tropospheric column data. The black solid and dashed lines indicate, respectively, the processor switch from 1.2.2 to 1.3.0, and the activation of the finer horizontal resolution. Stations are ordered by median tropospheric column.

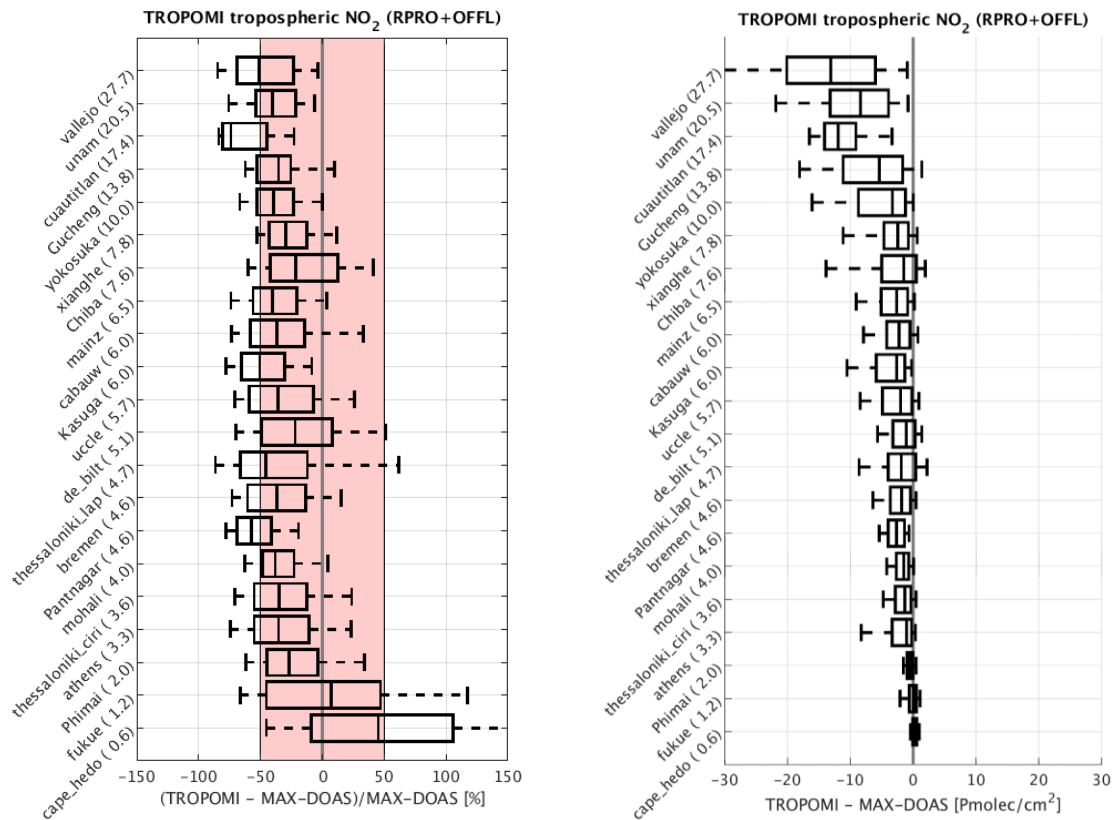


Figure 31: Box-and-whisker plots summarizing the bias and spread of the difference between S5p TROPOMI RPRO+OFFL and MAX-DOAS NO₂ tropospheric columns. The median difference is represented by a vertical solid line inside the box, which marks the 25 and 75% quantiles. The whiskers cover the 9-91% range of the differences. The shaded area represents the mission requirement 50% for the uncertainty. Values between brackets in the labels denote the median tropospheric column at the station. The time frame is from May 2018 until October 2020. This graph includes results for processor versions 1.1.0 and later. Stations are ordered by median tropospheric column.

When low tropospheric profiles are provided by the ground-based MAX-DOAS data partners (BIRA-IASB, ChibaU, JAMSTEC, and UNAM sites), 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 includes the 12 sites shown in **Figure 28** 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 is then generally decreasing (for 8 cases over 12 sites) by about 10 to 20% (see **Figure 28**). 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.

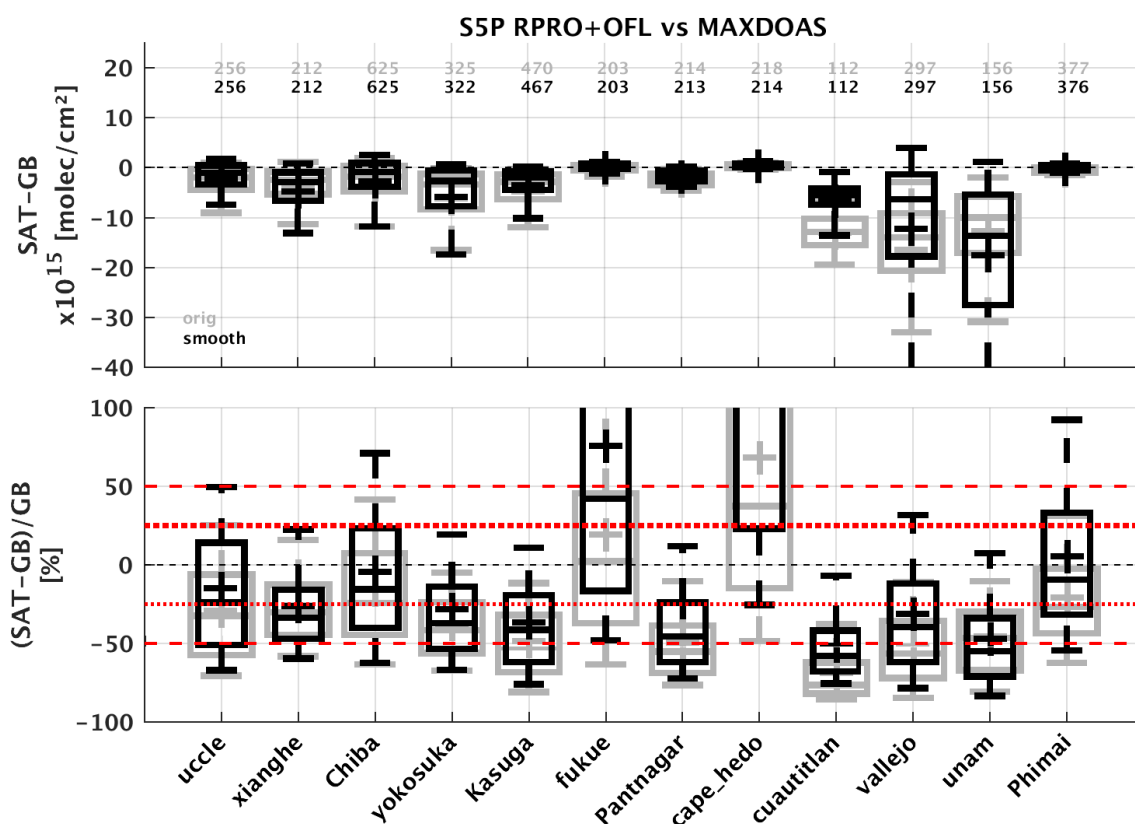


Figure 32: Box-and-whisker plots summarizing the bias and spread of the difference between S5p TROPOMI RPRO+OFFL and MAX-DOAS NO₂ tropospheric columns (grey) and the smoothed MAX-DOAS NO₂ tropospheric columns (black). The median difference is represented by a vertical solid line inside the box, which marks the 25 and 75% quantiles. The whiskers cover the 9-91% range of the differences and the crosses denotes the mean differences. The time frame is from May 2018 until November 2020.

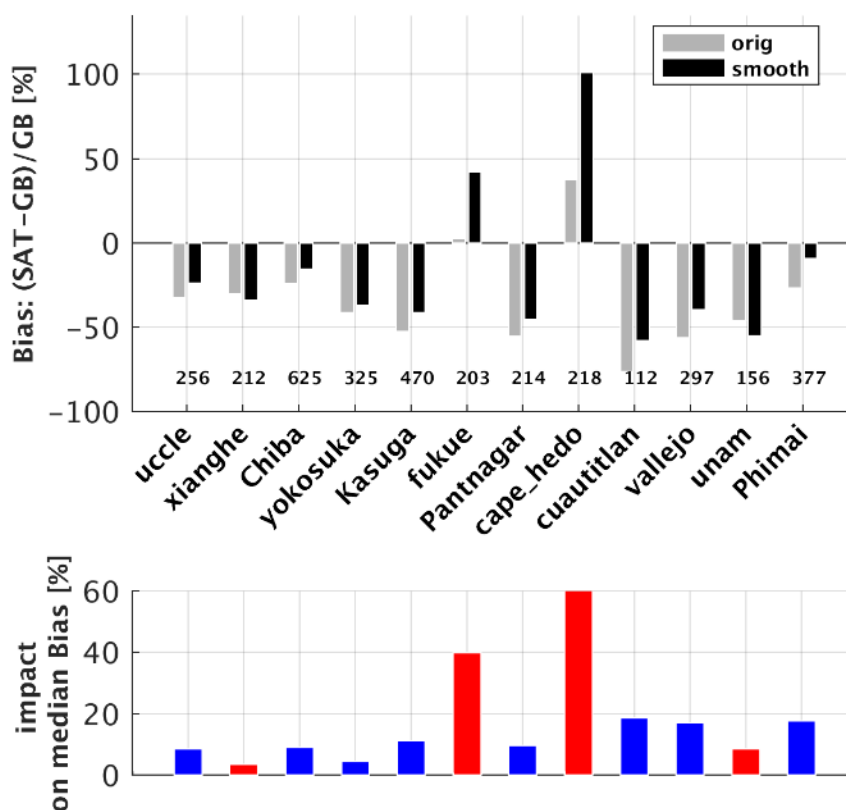


Figure 33: (upper panel) Bar plots summarizing the median bias of the difference between S5p TROPOMI RPRO+OFFL and MAX-DOAS NO₂ tropospheric columns (grey) and the smoothed MAX-DOAS NO₂ tropospheric columns (black). **(lower panel)** Bar plot summarizing the change on the median bias between the original 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 October 2020.

In summary, the tropospheric NO₂ bias depends on pollution level. On average, it is about -37%. It can be larger than -50% for extreme pollution and +7% in clean areas. Taking the sensitivities of the instruments into account by using the satellite averaging kernels with the sonde profiles, the bias can be reduced by up to 20%.

6.3.4.2 Dispersion

The median dispersion with respect to NIDFORVAL MAX-DOAS sites is 3.3 Pmolec/cm². Using again the three site types identified in the previous section (low polluted <2 Pmolec/cm², medium polluted from 3 to 14 Pmolec/cm², and extremely polluted >15 Pmolec/cm²), the dispersion (IP68/2) of the differences have median values of about 0.7 Pmolec/cm², 3.4 Pmolec/cm² and 12 Pmolec/cm², respectively. The uncertainty precision requirement of maximum 0.7 Pmolec/cm² is only satisfied for the clean-sites ensemble. It must be noted that also MAX-DOAS uncertainty sources and comparison errors contribute to the dispersion. Moreover, systematic errors (e.g., seasonal cycle) can contribute to the dispersion. A part of the systematic error component can be removed by calculating the spread around the OLS regression line instead of the standard deviation between S5P and MAX-DOAS data (Schneider et al., 2006). The residual spread at the different sites is approximately 2 Pmolec/cm². There is a reasonably good correlation between S5P tropospheric column and MAX-DOAS data, with the Pearson R varying between 0.50 (Uccle) and 0.83 (Athens).

6.3.4.3 Dependence on influence quantities

Two key influence quantities for observations of tropospheric NO₂ are aerosol optical depth (AOD) and satellite cloud (radiance) fraction (CRF). The first one is retrieved within the MAX-DOAS NO₂ analysis. No clear dependence of the bias on either property is seen from comparisons with MAXDOAS stations in **Figure 28**.

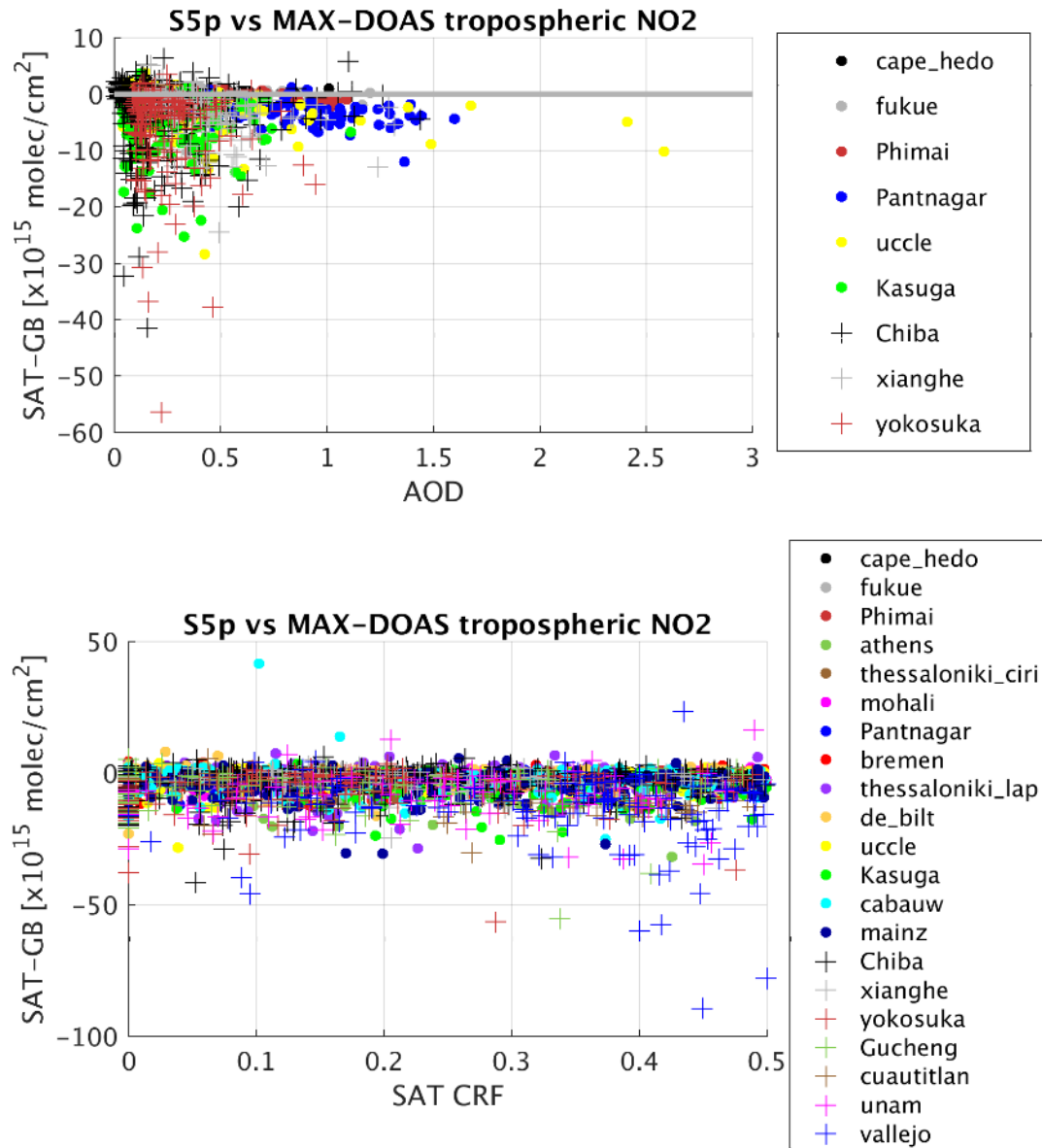


Figure 34: Differences between S5p TROPOMI RPRO+OFFL and MAX-DOAS NO₂ tropospheric columns as a function (upper panel) of MAX-DOAS AOD and (lower panel) S5p cloud radiance fraction (CRF). The time frame is from May 2018 until October 2020.

6.3.4.4 Seasonal and shorter term variability

Figure 35 presents the seasonal cycle of the difference between S5P RPRO+OFFL and NDACC MAX-DOAS tropospheric NO₂ (all comparison pairs reported on a single year). It should be noted that the RPRO is version 01.02.02 for the period of May-October 2018. A closer look at the different years and thus different versions do not show any impact of different versions on the validation.

S5P reports lower values than MAX-DOAS in late fall and in wintertime, when tropospheric NO₂ reaches its largest abundance. Over the entire year the 30-day rolling median relative difference is within the mission requirements for the bias.

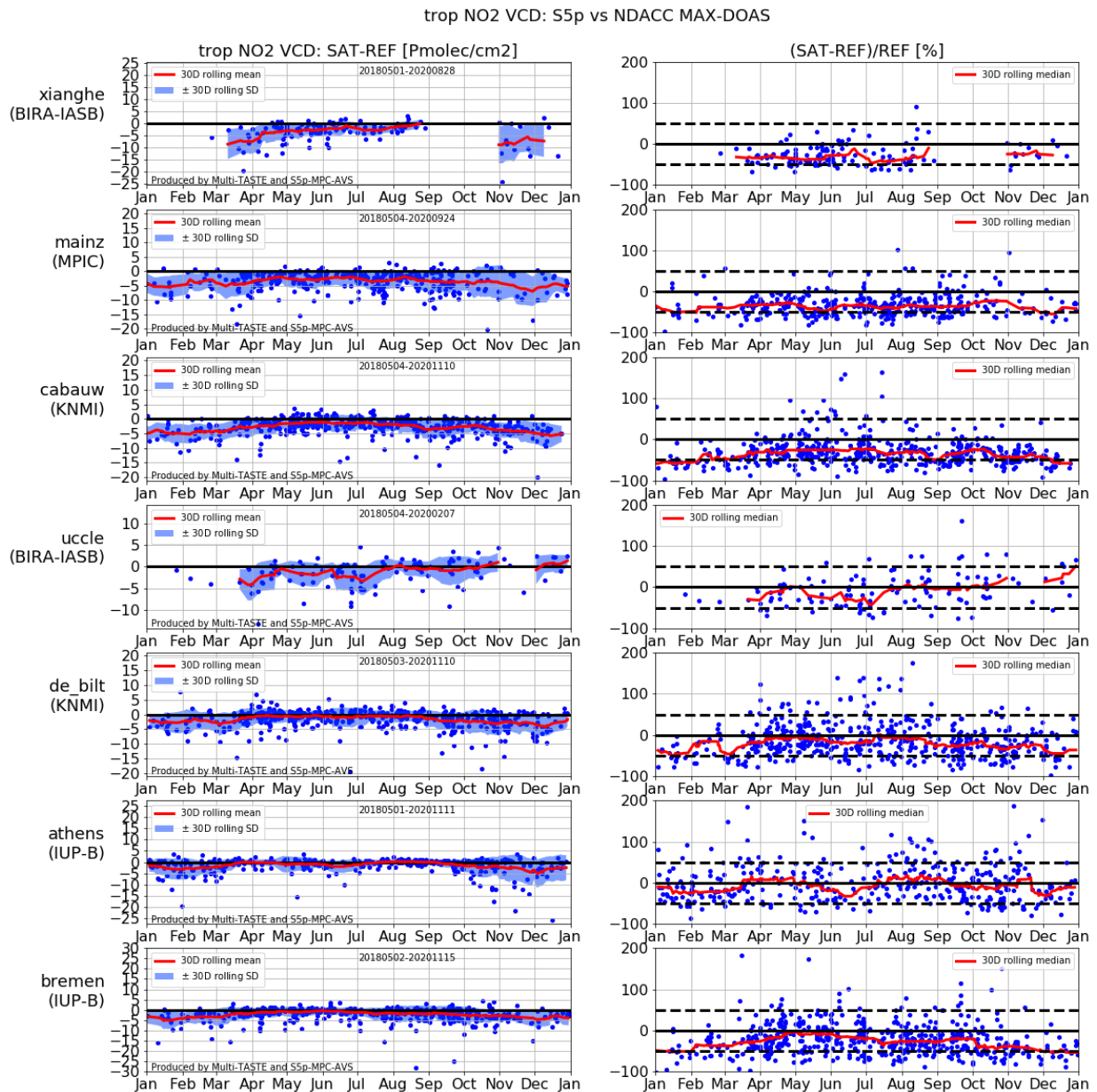


Figure 35: Seasonal cycle (with data mapped to one generic year) of the difference between S5P RPRO+OFFL and MAX-DOAS NO₂ tropospheric column data at six European stations. Difference (left) and relative difference (right). Data was obtained from the VDAF Automated Validation Server on 2020/11/25.

Comparisons with OMI show seasonal differences in the tropospheric columns. Three regions are selected: Europe (**Figure 36**), Eastern China (**Figure 37**, top), and as a reference sector the Pacific Ocean (**Figure 37**, bottom). The geometric column density (GCD) panels show that the TROPOMI slant column retrieval gives on average slightly larger NO₂ slant columns, with differences of a few percent. These differences are most likely related to details of the settings of the DOAS retrieval. For polluted areas the tropospheric column of TROPOMI is lower than OMI's, with differences varying from a few to -40%; the largest differences occurring in winter. In the clean Pacific Ocean area, tropospheric columns are smaller by a factor of up to 10. Here, TROPOMI is displaying higher values by about 20% in comparison to OMI.

Even though the comparison is done on the basis of equal sized grid cells, representativeness remains a major issue. Since OMI's ground pixels are much larger than TROPOMI's, in particular near the edges of the swath, the ground pixel capture different amounts of NO₂ and of cloud cover. In addition, OMI data suffers from the row anomaly.

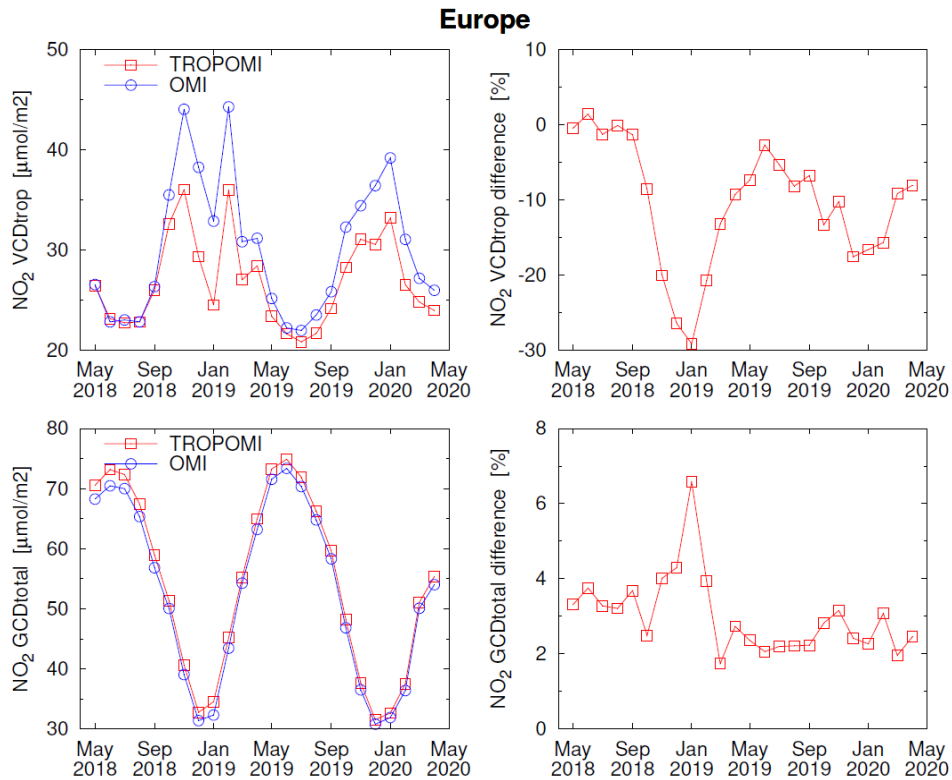


Figure 36: TROPOMI/OMI time evolution for Europe [10W,25E][35N,60N] of the tropospheric NO₂ column (VCDtrop, top rows) and the total geometric column (GCD, bottom rows) in absolute numbers (left columns) and as relative differences. TROPOMI data for pixels with qa_value > 0.75 and OMI/QA4ECV data with 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 averaged over calendar months. The total geometric column is the total slant column divided by the geometric air-mass factor: $GCD = SCD / AMF_{geo}$. The relative difference is defined as w.r.t. the average between the two.

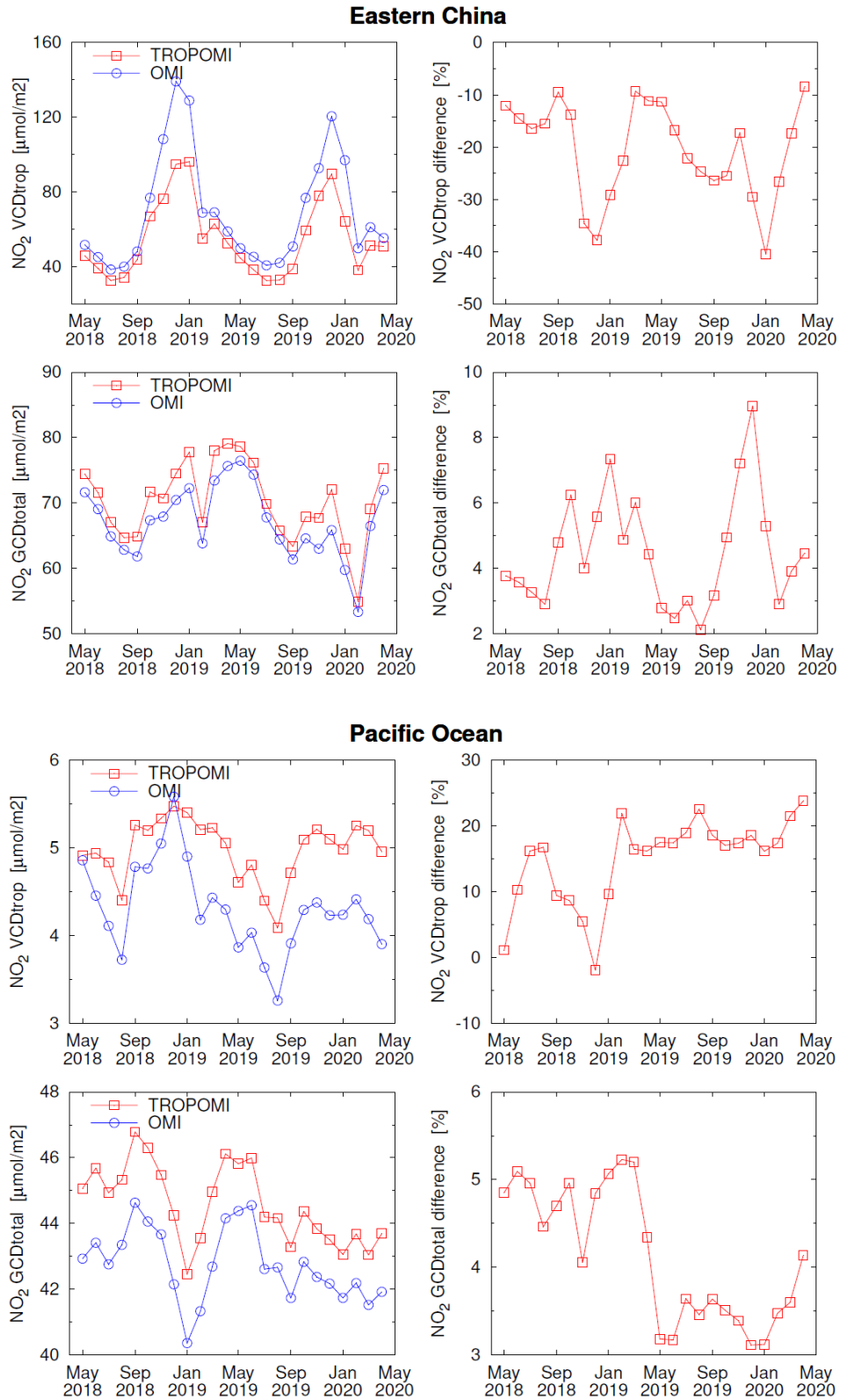


Figure 37: TROPOMI/OMI time evolution for Eastern China [110E,124E][21N,43N] and the Pacific region [170W,130E][60S,60N].

6.3.4.5 Geographical patterns

In general, no geographical patterns or artefacts can be detected in the latest L2_NO2 OFFL versions, as shown in the 6-month mean over central Europe (**Figure 38**).

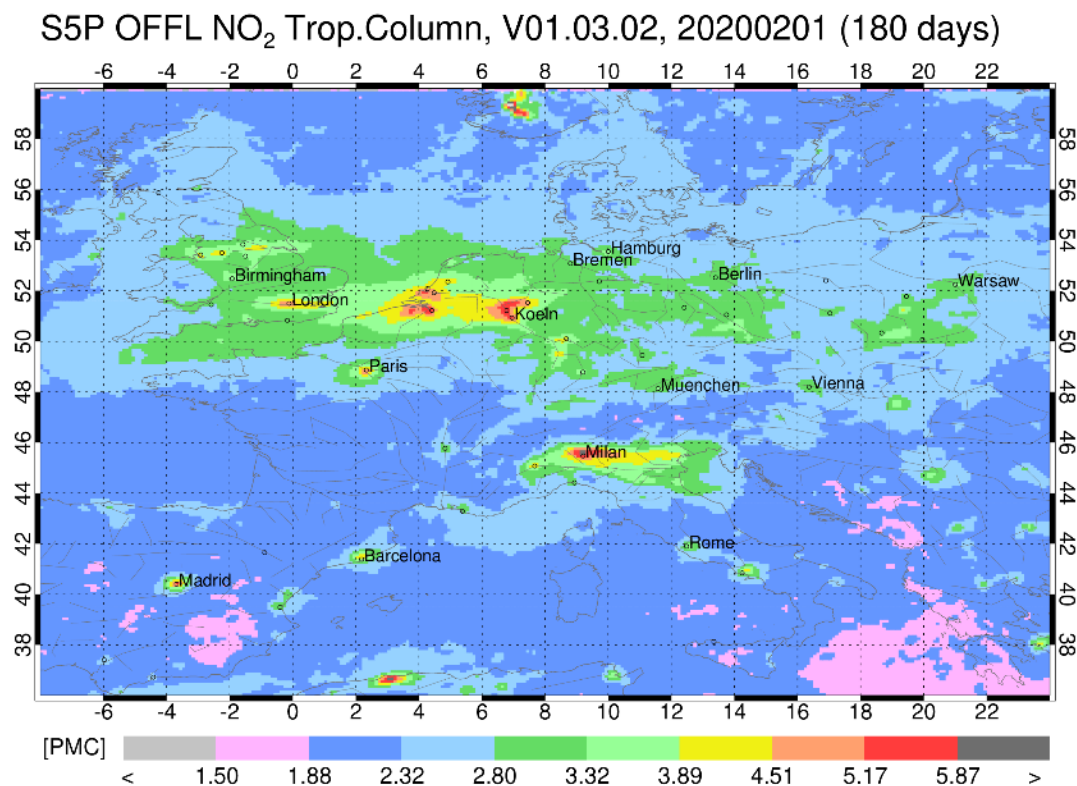


Figure 38: S5P OFFL tropospheric NO₂ over central Europe. The data was binned on a 0.5°x0.5° grid for a 6-month period starting February 2020 using processor versions 01.03.02. A $qa_value > 0.5$ and a cloud fraction $CF < 0.6$ were chosen to reduce the amount of data and exclude cloudy scenes. 1 PMC means 1 Pmole/cm².

6.3.4.6 Switch to smaller ground pixel resolution

No issue to report. Total column data show no effect as shown in Subsection 6.3.5.6.

6.3.4.7 Other features

Ordinary linear regression (OLS) of S5P (y) vs MAX-DOAS (x) yields fairly good correlation coefficients (0.58-0.83, see before) but low slopes. In $S5P = a \cdot MXD + b$, a varies between 0.3 (Bremen, Uccle) to 0.5 (Athens). It is known however that this approach is correct only in the limit that all random error is in y. OLS of MAX-DOAS vs S5P (i.e., assuming the opposite limit that all random error is in x), one obtains slopes closer to unity: in $S5P = a \cdot MXD + b$, a varies now between 0.72 (Athens, Mainz) and 1.2 (Uccle).

6.3.5 Total NO₂ column

6.3.5.1 Bias

Based on measurements from 26 Pandora stations available at VDAF-AVS (**Figure 397**) between 64.9°N and -45.8°S, the median bias (9783 co-locations, 2020-11-25) is -5.6% 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, the bias is +4.5% for the 16 less polluted stations (<7 Pmole/cm²) and -23.7% for the 10 high polluted stations.

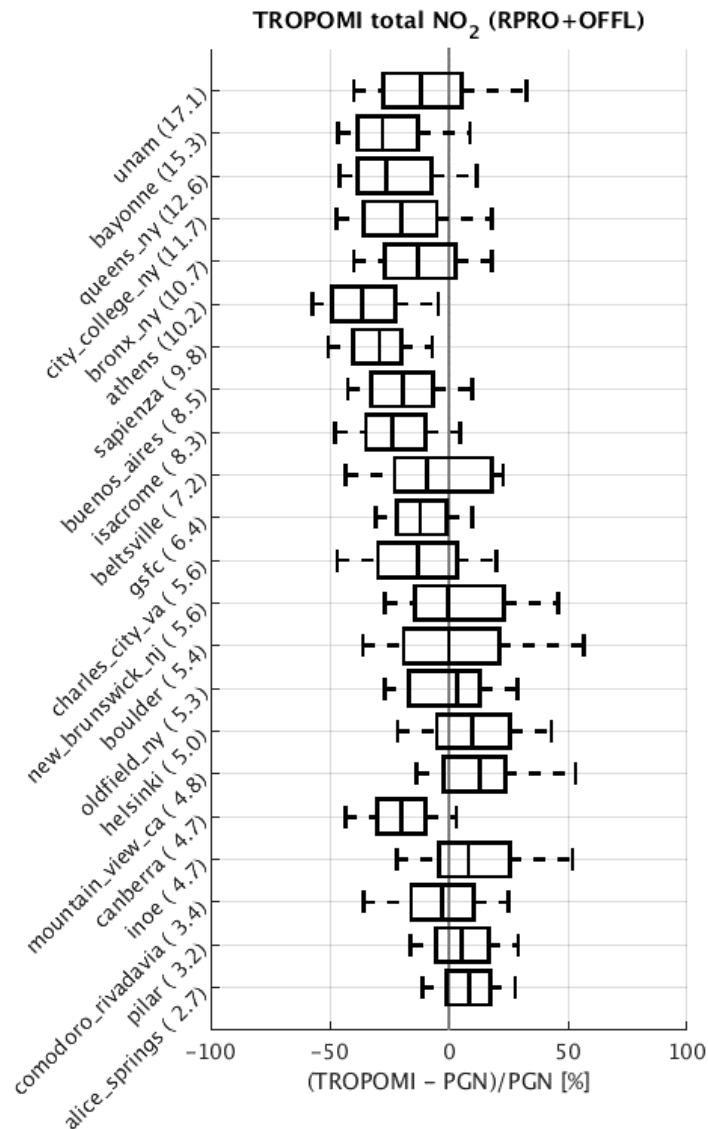


Figure 39: Box-and-whisker plots summarizing the relative bias and spread [%] between S5p TROPOMI RPRO+OFFL (processor version 1.2.2 and later) and PGN Pandora NO₂ total column data. Conventions of the boxplots are identical to Fig. 21. Stations are ordered by median total column. The time frame is May 2018 to October 2020.

We highlight 3 different comparison cases here. At Alice Springs (Australia), where the total NO₂ column values are mostly between 2-4 Pmolec/cm², a small positive mean bias of 0.2 Pmolec/cm² is seen (+8% median bias). A wider distribution of NO₂ values (2-30 Pmolec/cm²) is found at Bronx (New York, United States) where the bias is -3 Pmolec/cm² (-15%). Finally, at Sapienza (Rome, Italy) where column values can reach up to 40 Pmolec/cm², the bias is highest with -7 Pmolec/cm² (-46%). Locally enhanced NO₂ probably contributes to this.

6.3.5.2 Dispersion

The dispersion of the S5P and PGN Pandora differences depends strongly on the station. Small uncertainties (standard deviations) are observed at Alice Springs, Izaña and Mauna Loa (0.5-0.7 Pmolec/cm²) that are within the mission precision requirement, and higher values elsewhere (e.g., 4-6 Pmolec/cm² at New York Bronx and at Sapienza Rome, 7 Pmolec/cm² at New York City College). The site averaged dispersion is 2.9±2.1 Pmolec/cm². The Pearson-R varies from relatively low (e.g., 0.3 at Charles City) to high (0.89 at New York Bronx).

6.3.5.3 Dependence on influence quantities

None to report.

6.3.5.4 Short term variability

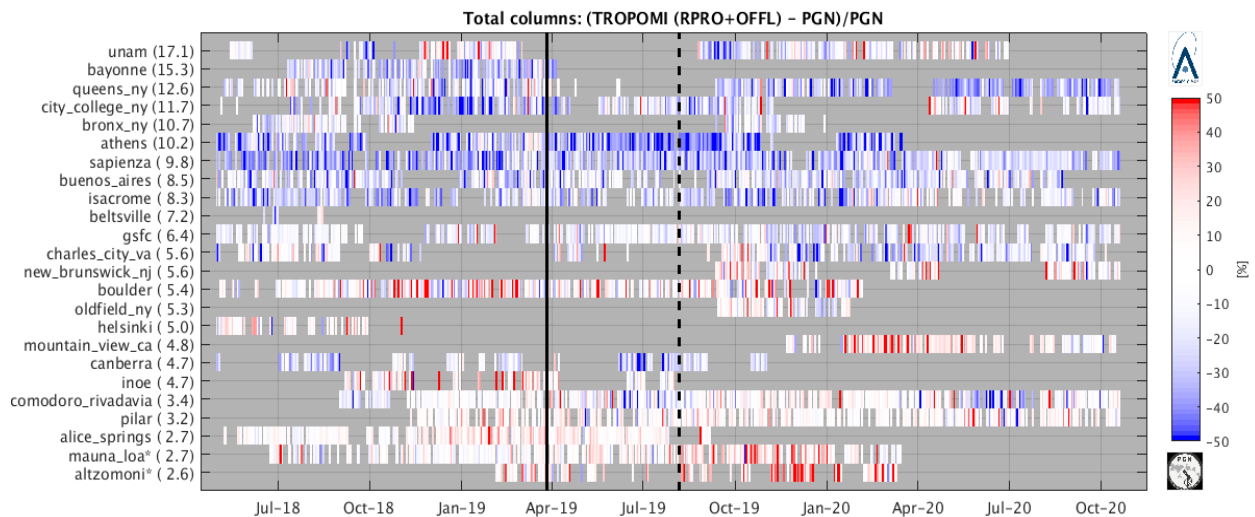
None to report.

6.3.5.5 Geographical patterns

None to report.

6.3.5.6 Processor version change and switch to smaller ground pixel resolution

A major processor version change (1.2 to 1.3) occurred for OFFL L2_NO2 at orbit 7425 on 20 March 2019. On 6 August 2019, there was a change in the TROPOMI ground pixel size. Both changes were investigated by having a close look at the S5P and PGN Pandora time series at individual stations. **Figure 40** shows that the bias and scatter of the difference are not affected by the pixel size change.



Update of Verhoelst et al., AMTD 2020–119, 2020, produced by BIRA–IASB for the S5p MPC VDAF on 02–Nov–2020 using NDACC reference data

Figure 40: Time series – from May 2018 until October 2020 – of S5P RPRO+OFFL and PGN Pandora NO₂ total column differences [%]. 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. Stations are ordered by median total column.

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 Stratospheric NO₂ 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 August 12, 2020. It reveals differences mostly below the mission requirement on precision (0.5 Pmolec/cm²). The RMSD is 0.2 Pmolec/cm², with values up to 0.5 Pmolec/cm². Some features are due to a different stratosphere/troposphere separation (e.g. over Greenland, positive difference in stratosphere, negative in troposphere, **Figure 41**), while other are related to differences in total columns, here in particular at the orbit start and end. 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.

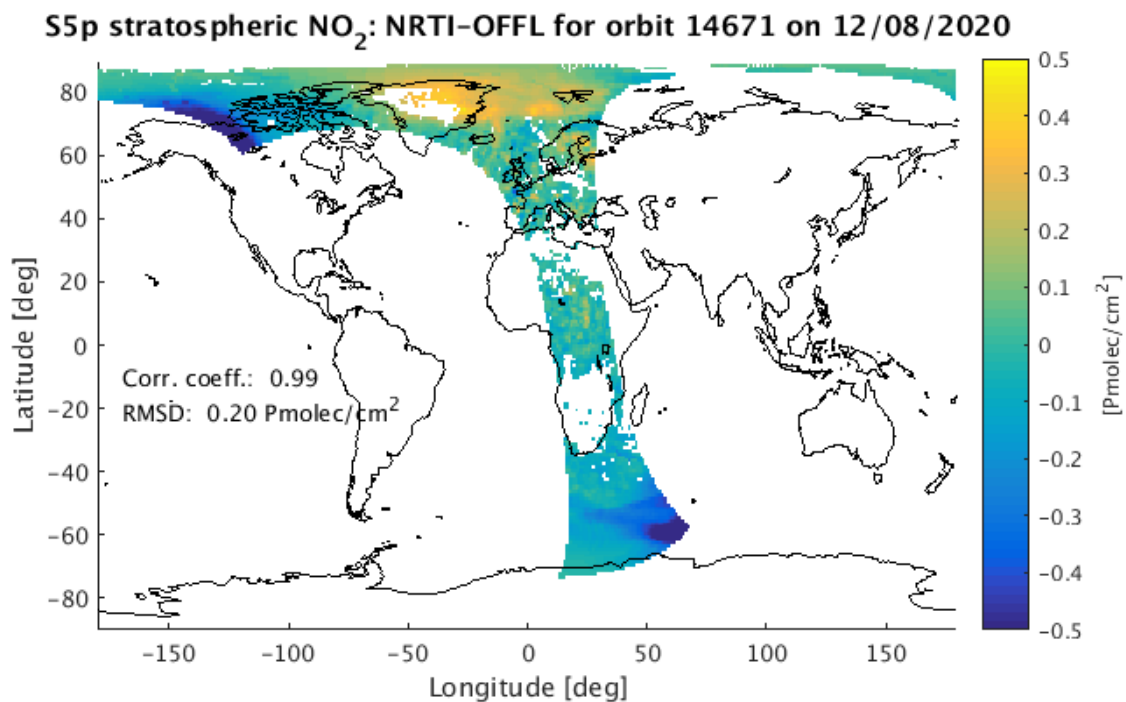


Figure 41: Difference between S5P NRTI and OFFL stratospheric NO₂ column data for a single orbit (gridded to 1°x1° resolution).

A comparison of NRTI and OFFL pole-to-pole validation graphs is shown in **Figure 42**, illustrating that in the end, NRTI performs very similarly to OFFL with a bias of -0.2 Pmolec/cm².

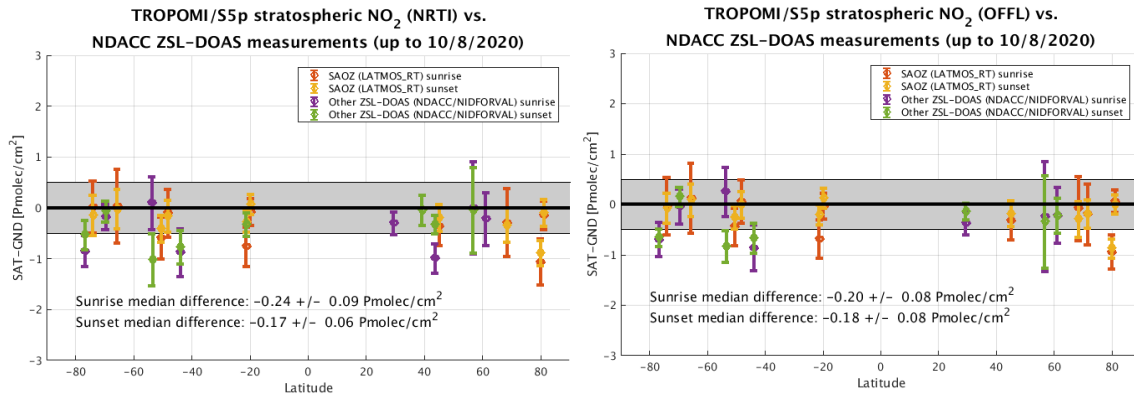


Figure 42: Meridian dependence of the mean (the circular markers) and dispersion ($\pm 1\sigma$ error bars) of the differences between S5p TROPOMI L2_NO2 (NRTI in the upper panel, OFFL in the lower) stratospheric column data and ZSL-DOAS reference data, represented at individual stations from the Antarctic 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 time period of phase E2 until August 2020.

6.4.2 Tropospheric NO2 Column NRTI vs OFFL

To demonstrate the closeness of L2_NO2 NRTI and OFFL products at the available MAX-DOAS sites, L2_NO2 NRTI (processor version 01.00.02 to 01.03.02) and L2_NO2 OFFL (RPRO processor version 01.02.02 + OFFL processor version 01.02.00 to 01.03.02), 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 the following **Table 5**. Similar conclusions on the closeness of NRTI and OFFL can also be drawn for the other sites.

Table 5 – 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: 382 common co-locations Orbits range from 3364 (2018-06-07) to 16019 (2020-11-15).			
	NRTI-OFFL	NRTI-MXD	OFFL-MXD
Mean(diff) $\pm$ sem*	0.08	-2.11 $\pm$ 0.18	-2.19 $\pm$ 0.18
Median(diff)*	0.06	-1.26	-1.28
Std(diff)*	0.4	3.5	3.5
1/2 IP68(diff)*	0.2	2.4	2.4
Pearson R	0.98	0.54	0.54
Slope	1.01	0.25	0.25
Athens: 342 common co-locations Orbits range from 3392 (2018-06-09) to 15962 (2020-11-11).			
	NRTI-OFFL	NRTI-MXD	OFFL-MXD
Mean(diff) $\pm$ sem*	0.11	-1.33 $\pm$ 0.19	-1.44 $\pm$ 0.19
Median(diff)*	0.09	-0.15	-0.23
Std(diff)*	0.2	3.6	3.6
1/2 IP68(diff)*	0.1	2.2	2.2
Pearson R	1.00	0.82	0.83
Slope	1.02	0.50	0.49

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 value and the near unity slope of the linear regression. The daily relative difference is 7.2% with RMSD of 0.32 and Pearson R of 0.93 on 2020-10-01.

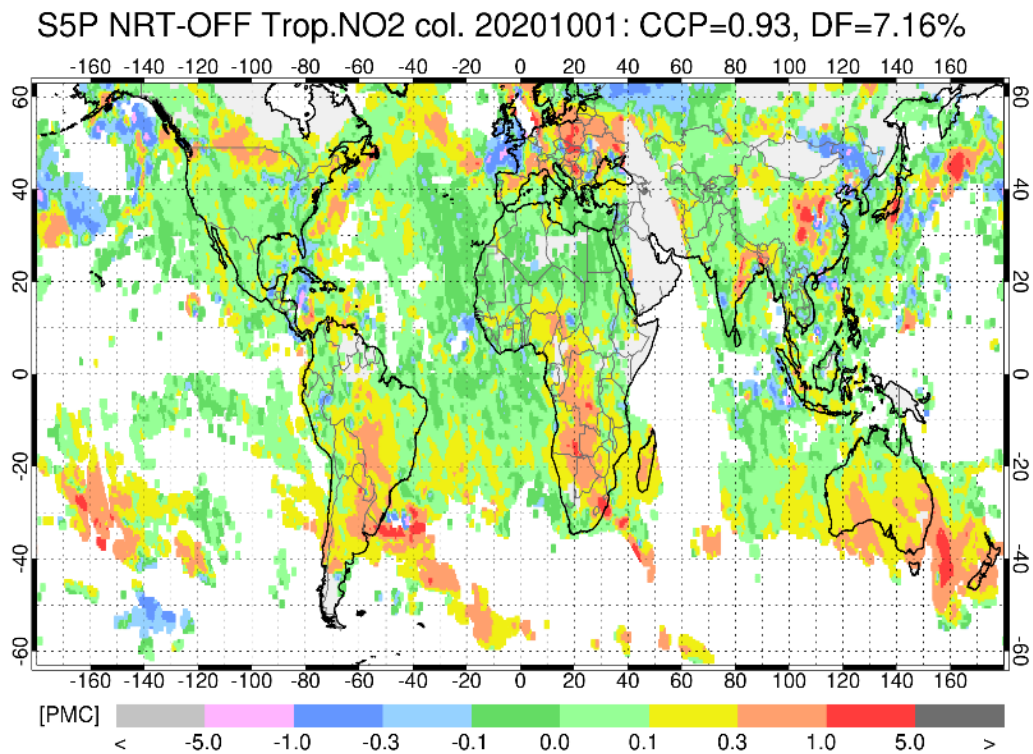


Figure 43: Difference between S5P NRTI and OFFL tropospheric NO₂ column data for one day (gridded to 0.5°x0.5° resolution). The difference is only calculated for columns above 0.5 Pmolec/cm².

6.4.3 Total NO₂ Column NRTI vs OFFL

The comparison of total NRTI vs. OFFL data show that both the overall values and the standard deviations are very close to each other (**Figure 44**). The global relative difference is in the range of 0.8%, with NRTI being slightly higher. After the change to PDGS 2.7.0, the positive bias has changed within a few months to a negative bias. A single orbit comparison as previously analyzed for the stratospheric and tropospheric columns reveals a combination of the features already seen in these subcolumns (**Figure 45**).

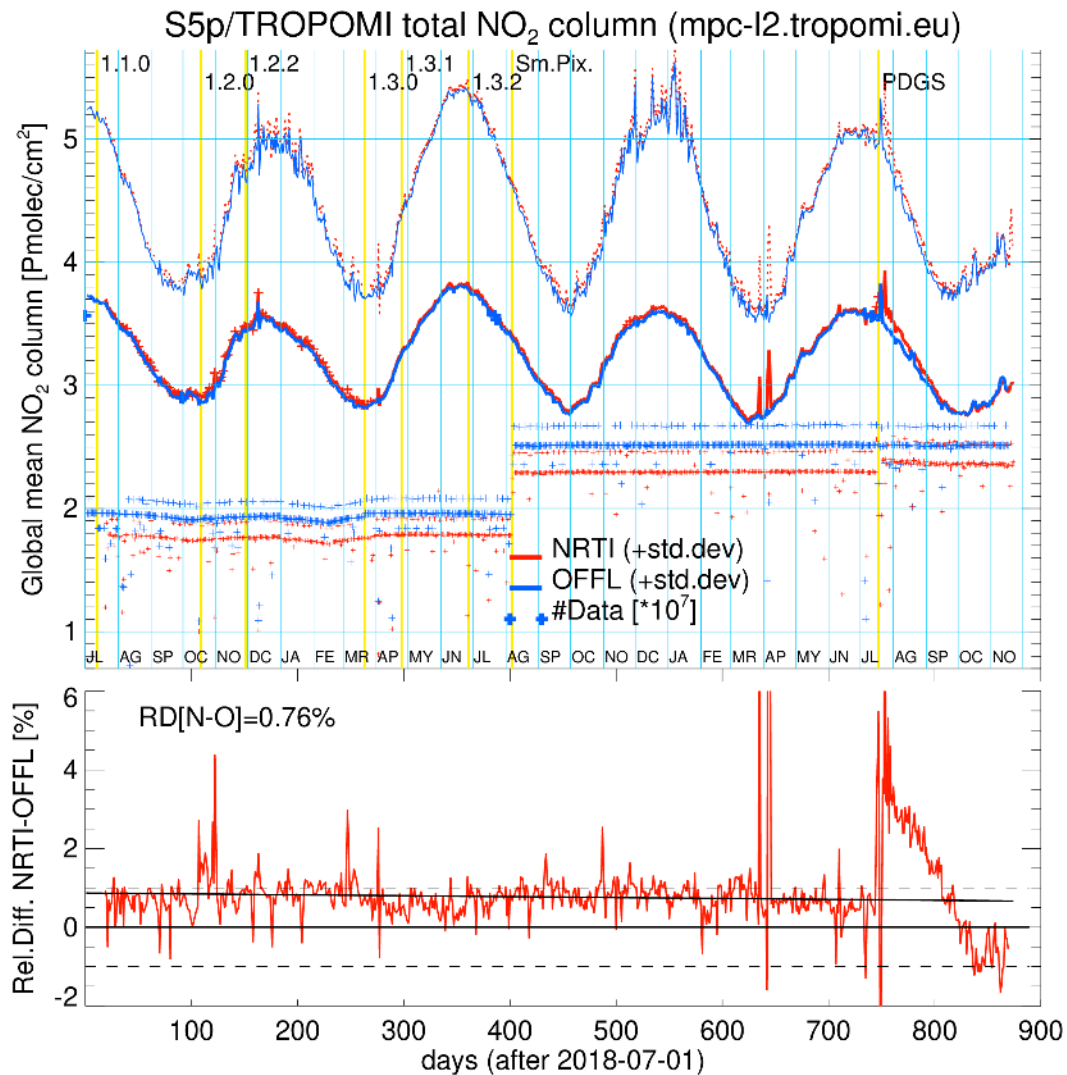


Figure 44: Time series for phase E2 (July 2018 – November 2020) of the global mean of NRTI (red) and OFFL (blue) NO₂ total column data [Pmolec/cm²]. The added 1 σ standard deviations are shown as thin lines and the number of NRTI/OFFL data points as crosses with values divided by 10⁷. Superimposed are the points of processor changes, small pixel switch, and latest PDGS switch to 2.7.0 as yellow vertical lines. The lower plot shows the difference [%] between OFFL vs. NRTI daily means. Data is taken from the TROPOMI QC portal.

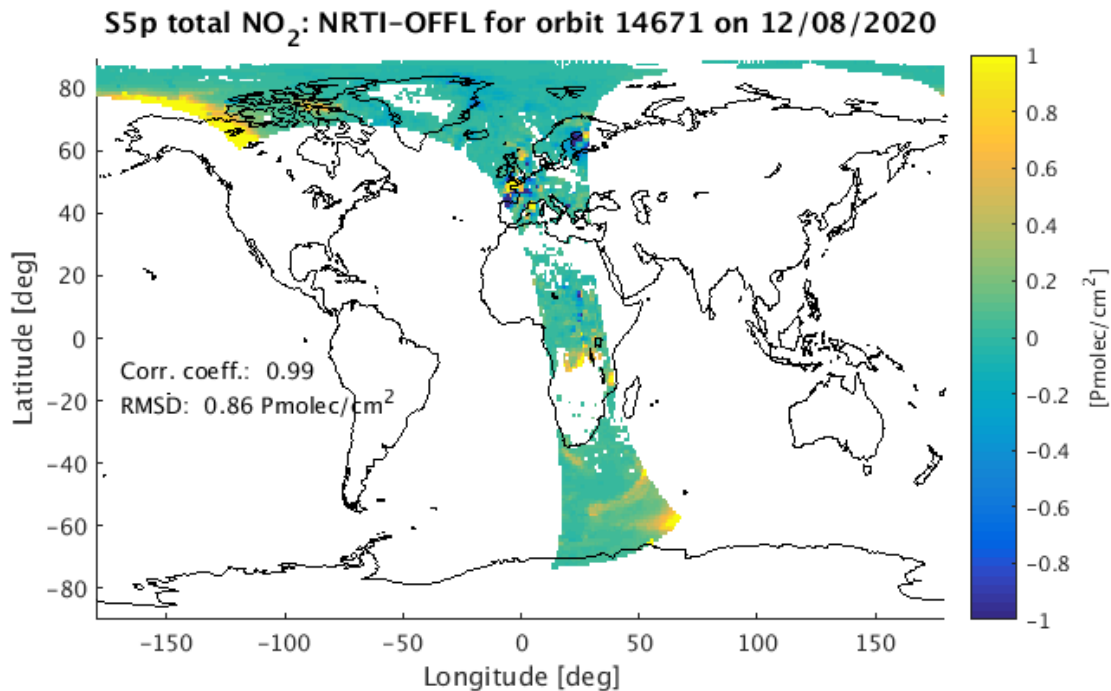


Figure 45: Difference between S5P NRTI and OFFL total NO₂ column data for a single orbit (gridded to 1°x1° resolution).

6.5 Internal consistency of the validation results

This section focuses on the internal consistency checks of the NO₂ validation results. This rely on analysis of potential source of inconsistencies (such as impact of different NO₂ cross-sections) or synergistic analysis of several instruments types.

6.5.1 NO₂ cross-sections:

A potential source of inconsistencies between the different data products lies in the NO₂ cross sections that are used. An overview of the different NO₂ cross sections choices made for each instrument is provided in **Table 6** (adapted from Verhoelst et al. (2020)). 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 be overestimated by up to 10% when the column is mostly stratospheric or underestimated 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 6. NO₂ cross section source and temperature for the different instruments (Verhoelst et al., 2020).

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	NIWA instruments
ZSL-DOAS	Harder et al. (1997)	227K	
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 (2020) 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 NO₂ columns, respectively. By subtracting the TROPOMI stratospheric VCD from the direct-sun total columns, an estimation of the tropospheric NO₂ can be obtained and compared to MAX-DOAS results (Pinardi et al., 2020). Current sites with these 2 types of instruments are Athens, Thessaloniki, Xianghe, Uccle, and Unam.

Preliminary comparisons have been drawn, 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 NO₂ 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. The L2_HCHO RPRO data product version 01.01.05 was released in May 2019, covering the period from 26 June to November 2018. RPRO 01.01.02 covered the period before. The latest L2_HCHO NRTI and OFFL product version 02.01.03 was released in July 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

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

The comparison methodology is described 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 a priori profile of the FTIR is substituted with that of S5P TROPOMI L2_HCHO, yielding a corrected FTIR profile. (ii) This corrected FTIR profile is smoothed with the S5P averaging kernel (Rodgers and Connor, 2003). (iii) Scaling is applied to take into account altitude differences between pixel level and station altitude. (iv) The individual manipulated FTIR columns are averaged, and the individual S5P manipulated pixel columns are averaged.
- The relative bias at a single station is estimated by the median relative difference: $\text{Med}[(\text{SAT}-\text{REF})/\text{REF}]$. Absolute-scale dispersion is estimated by the scaled median absolute deviation from the median (MAD): $1.4826 \cdot \text{med}[(\text{abs}(\text{DIFF}-\text{med}(\text{DIFF})))]$. 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

S5P 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% and a precision better than 30%. The vertical sensitivity of MAX-DOAS is different from that of S5P HCHO. MAX-DOAS has a higher sensitivity close to the surface and a lower sensitivity at higher altitude, while 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:

1. **VDAF Automated Validation server:** S5p pixel are kept for `qa_value` above 0.5, 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.
2. **NIDFORVAL AO project:** S5P pixels with a `qa_value` above 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.

7.2.2 Satellites

S5P 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). Two versions of the OMI L2 HCHO data product are considered: (1) the NASA L2 product (10.5067/Aura/OMI/DATA2015), and (2) the QA4ECV L2 product (<http://doi.org/10.18758/71021031>). The first has the advantage of being operational and completely independent from TROPOMI retrievals. The second offers the advantage to be produced by the same European consortium as the TROPOMI product and thus results can directly be compared because the algorithms have been made as consistent as possible.

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` above 0.5 (no error flag, cloud radiance fraction at 340 nm < 0.5, SZA $\leq 70^\circ$, surface albedo ≤ 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). It 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 streams is as follows:

- Comparisons of S5P HCHO with NDACC FTIR follow the methodology of Vigouroux et al (2020). The current number of stations is 27 (extended with 2 new sites: Karlsruhe, Germany; and Hefei, China) and covering the period May 2018-May 2020.
- 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 November 2020.
- Data from nine stations are used in the NIDFORVAL AO project. The period covers May 2018 to May-July 2020, depending on the station.

7.3.3 Bias

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

7.3.3.1 Fourier Transform Infrared Spectrometers

HCHO has a positive bias of $+20 \pm 5\%$ for clean sites (< 2.5 Pmolec/cm²) and a negative bias of $-29 \pm 1\%$ for high emission sites (HCHO > 8 Pmolec/cm²) in comparison to correlative data from 27 NDACC FTIR stations, covering the period from May 2018 to May 2020. This is illustrated in **Figure 46**. 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 0.97 ± 0.03 Pmolec/cm² and a proportional bias of 0.65 ± 0.03 is obtained (updated from Fig. 4 in Vigouroux et al., 2020).

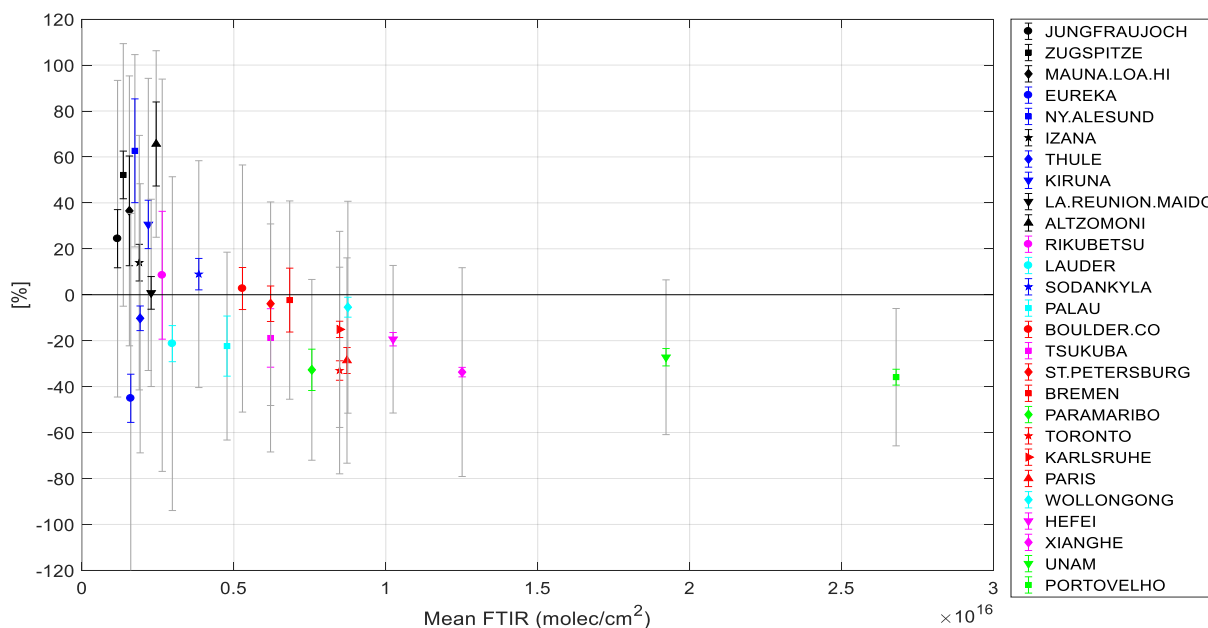


Figure 46: Bias at each station (in %) as a function of the mean FTIR total columns (molec/cm²). The gray bars are the systematic uncertainty on the differences, and the colored error bars are the 2- σ error on the bias. Updated from Fig. 3 of Vigouroux et al., 2020; temporal range here is May 2018 to May 2020.

7.3.3.2 MAX-DOAS UV-visible Spectrometers

The S5P L2_HCHO median bias is -21% with a larger scatter of 35% with respect to MAX-DOAS total HCHO data from 5 VDAF-AVS different stations for 2398 collocations until November 2020. The median difference is negative by up to -5.4 Pmolec/cm² at 4 sites and only positive with 0.7 Pmolec/cm² in Uccle. In the NIDFORVAL AO project 4 more stations are available. Then we get a larger median bias of -38%. These results are within the mission requirement of 80% for the bias. The biases at each station are provided in **Figure 47** (upper panel) as a function of the mean DOAS levels for the period May 2018 to July 2020.

The bias improves at most of the stations (see **Figure 47**, lower panel) when taking the difference in a-priori profiles and vertical resolution of the instruments into account (Rodgers and Connor, 2003). Using the averaging kernels from 5 stations (**Figure 48**), the smoothed (DOAS_{smoo}) median bias is -22%. This technique cannot be applied for stations not providing averaging kernels (Cabauw, De Bilt, Thessaloniki, and Mainz).

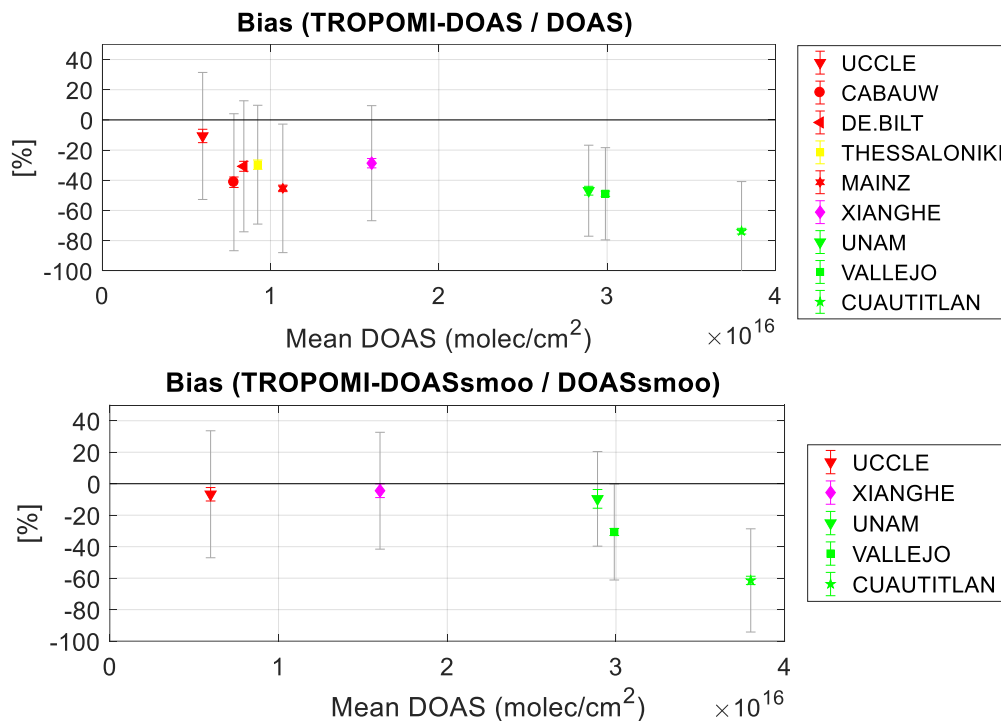


Figure 47: Bias at each station (in %) as a function of the mean DOAS total columns (10¹⁶ molec/cm²). The gray bars are the systematic uncertainty on the differences, and the colored error bars are the 2-σ error on the bias. Top panel: DOAS is the normal product without any modification. Bottom panel: DOAS data after Rodgers and Connor (2003) is applied (a priori substitution and smoothing with the TROPOMI averaging kernels).

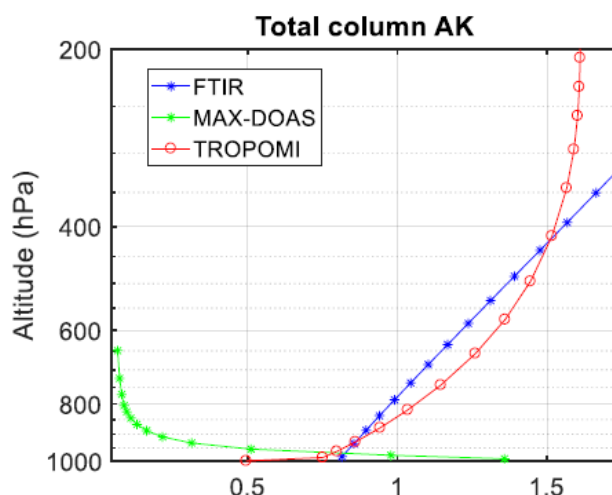


Figure 48: Typical total column averaging kernels for S5P TROPOMI and the two ground-based instrument types: FTIR (blue), MAX-DOAS (green), and TROPOMI (red). This illustrates the problem of the vertical smoothing difference error in these comparisons, as the instruments “see” different parts of the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

Further separating the biases for low (<2.5 Pmolec/cm²) and high (>8 Pmolec/cm²) HCHO levels as done for FTIR comparisons, we get for the DOAS_{smoo} comparisons $+9\pm25\%$ and $-28\pm2\%$, respectively. The bias for high-HCHO level is in very good agreement with FTIR results. 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**).

On the contrary, if we use the scatter plots to derive the slope and intercept as done with FTIR, we obtain 3.13 ± 0.22 Pmolec/cm² and 0.49 ± 0.03 for the constant and proportional biases, respectively. This is not in agreement with FTIR. The much larger proportional bias obtained with DOAS_{smoo} is mainly due to the Cuautitlan station. Excluding this station, we obtain a 0.61 ± 0.04 proportional bias in agreement with FTIR. The bias is then $-22\pm2\%$ for high HCHO which is smaller than FTIR.

In summary, the bias of S5P HCHO vs MAX-DOAS reduces from $\sim\sim 40\%$ to $\sim\sim 20\%$ when smoothing is applied to the MAX-DOAS profile. The bias becomes more negative for larger columns in both comparisons with FTIR and MAX-DOAS. With smoothing applied, the bias of S5P HCHO vs FTIR is $+20\pm5\%$ (<2.5 Pmolec/cm²) and $-29\pm1\%$ (>8 Pmolec/cm²). This is broadly consistent with S5P HCHO vs MAX-DOAS of $+9\pm25\%$ (<2.5 Pmolec/cm²) and $-28\pm2\%$ (>8 Pmolec/cm²). But it should be remarked that the number of MAX-DOAS stations where smoothing can be applied is small and moreover clean sites are not well represented.

7.3.3.3 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 15 years (2005-2019), including two 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.

While the OMI product is based on the O₂–O₂ absorption feature around 477 nm, and consider a fixed cloud albedo of 0.8 (Veefkind et al., 2016), the TROPOMI product uses the S5P operational cloud product in CRB mode (OCRA/ROCINN-CRB; Loyola et al., 2018). The S5P ROCINN algorithm is based on the O₂ A-band around 760 nm and retrieves both 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.

Figure 49 presents the mean bias between OMI and TROPOMI HCHO tropospheric columns for 25 regions. Numbers are provided for daily averaged columns applying a cloud correction (left) or not (right). Systematic differences due to the cloud correction up to 25% can be observed over Tropical regions where the clouds are the highest in altitude (Africa, South America, South Asia). A smaller but systematic effect (10-15%) is also observed over mid-latitudes polluted regions such as China, India, US or Europe. It is therefore advised to use clear VCDs when comparing satellite datasets using different cloud products.

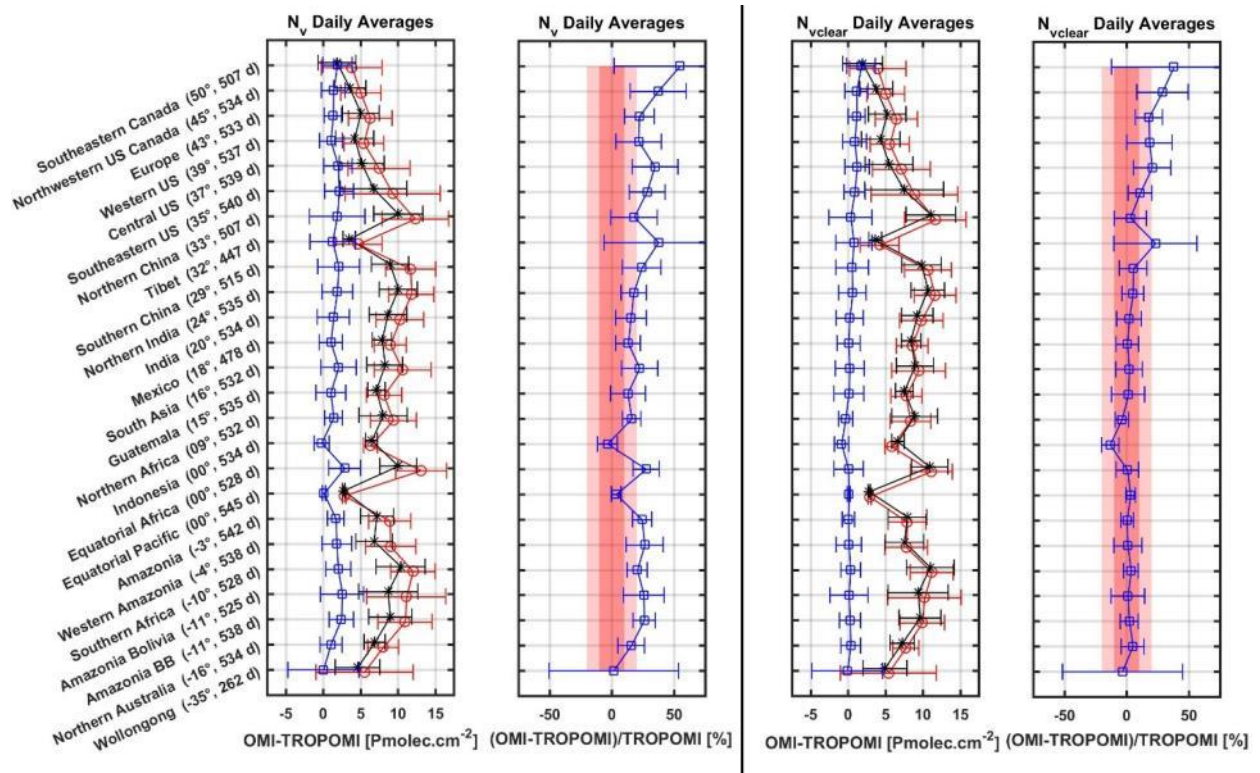


Figure 49: Absolute and relative bias between OMI and TROPOMI HCHO daily averaged columns using cloud corrected AMF (left) or clear-sky AMF (right) for a selection of regions. Regions are displayed by decreasing latitude. The mean OMI (red) and TROPOMI (black) columns are plotted together with the absolute bias.

When comparing clear VCD ($N_{v,clear}$), the bias between OMI and TROPOMI is below 10% in 18 over the 25 selected regions. Mid-latitudes/mid-emissions regions such Europe, Central US, Western US and North Western Canada present a remaining bias of about 15%. Higher than 45° in latitude, Southeastern Canada and North Western US-Canada regions present a larger bias of about 30%. The mean columns are lower over those regions, and differences in sampling (pixel size and OMI row degradation) start playing a bigger role.

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 $\text{SQRT}(\#\text{pixels})$. 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 27 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 (20km around the station). The MAD for the 27 stations taken together is 2.4 Pmolec/cm² corresponding to a theoretical single-pixel dispersion of $2.4 \times \text{SQRT}(\#\text{pixels}) = 13$ 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: 1.33 Pmolec/cm² for the 33-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 dispersion (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 Mainz. This is within the mission requirement of precision of 12 Pmolec/cm². Using the 20km-averaged-pixels within NIDFORVAL (~about 35 pixels here), as done for FTIR, we obtain a MAD of 3.6 and 3.8 Pmolec/cm² at Cabauw and De Bilt, respectively, when Rodgers and Connor (2003) is not applied as in the MPC. This corresponds to a single pixel precision of 21-23 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 too large collocation error, the MAD is 2.5 Pmolec/cm² for the 33-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_{smo} 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 $\text{SQRT}(\#\text{pixels})$, the dispersion requirement is now slightly (FTIR) and strongly (MAX-DOAS) higher than the dispersion requirement. But comparison error and FRM random error in the case of MAX-DOAS make an important contribution at polluted 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 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. If the observations are averaged to a common horizontal resolution, we expect a factor 5 improvement of precision from OMI to TROPOMI, thanks to an increase of at least a factor 15 in the sampling of TROPOMI. The median deviation of the differences between OMI and TROPOMI is about 2 Pmolec/cm². Larger values are observed over emission regions, indicating a larger variability. 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.

7.3.5 Dependence on influence quantities

None to report.

7.3.6 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 the time series at Karlsruhe and Hefei. In agreement with the biases discussed in Sect. 7.1.3, TROPOMI shows negative bias when the HCHO columns are the highest (during the July maximum). The correlation between the TROPOMI and FTIR monthly means at the 27 sites together is 0.91. **Figure 55** shows that the same seasonal variability is observable at other FTIR stations at Xianghe and Sodankylä.

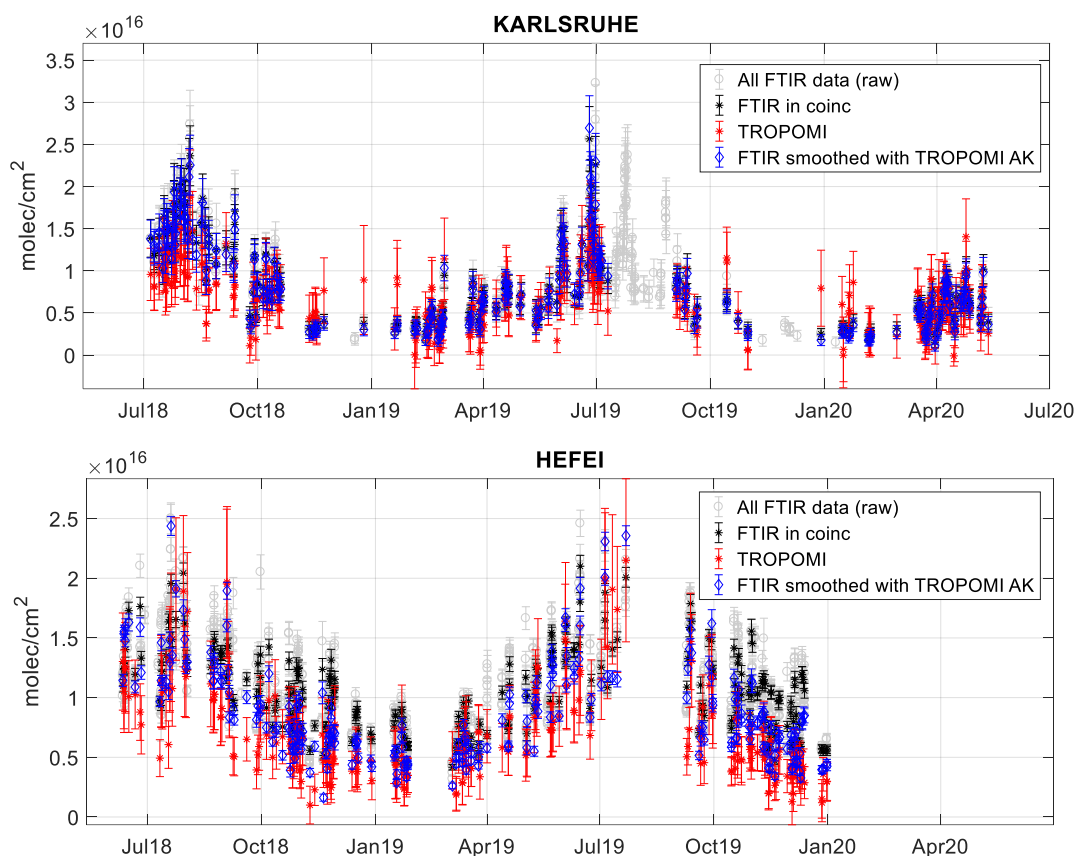


Figure 50: TROPOMI and FTIR HCHO time series at Karlsruhe and Hefei.

MAX-DOAS UV-Visible Spectrometers

The comparisons of TROPOMI and DOAS data within the NIDFORVAL project show a correlation of the monthly means of 0.82 for the 9 stations, and of 0.83 after smoothing for the 5 stations providing averaging kernels. But it varies quite a lot depending on the stations: from 0.97 for Xianghe down to -0.09 at UNAM (Mexico City). **Figure 51** shows time series at Uccle (monthly mean correlation=0.82) and Vallejo (monthly mean correlation=0.07). The poor correlation at the three Mexican sites might be due to weaker seasonal cycle there which should be investigated in the future.

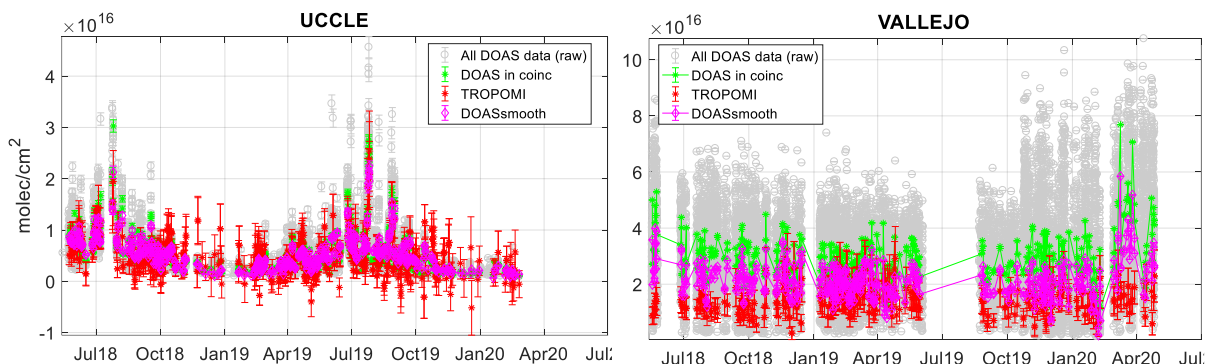


Figure 51: TROPOMI and DOAS time series at Uccle and Vallejo.

OMI QA4ECV comparisons

Day to day correlation between OMI and TROPOMI is very high above emission regions (**Figure 52**).

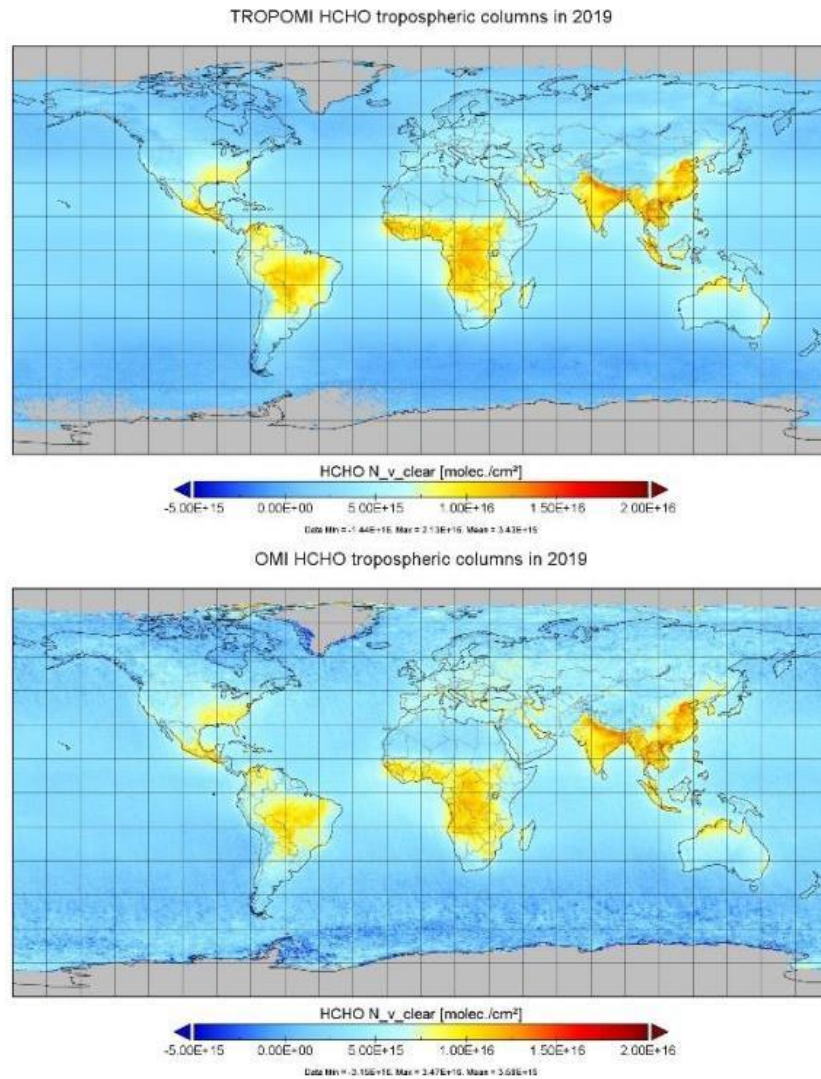


Figure 52: 2019 annual average of S5P TROPOMI (top) and Aura OMI (bottom) HCHO tropospheric column data (N_{v_clear}).

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 an overall very good agreement. Differences range from 2 Pmolec/cm² over Tropics to -2 Pmolec/cm² over mid-latitude regions. The gain in TROPOMI precision can be observed at the global scale, mainly at larger latitudes where the OMI sampling is the most affected in 2019.

The S5P_L2_HCHO data are seasonally averaged for spring (March-May 2018) and summer (June-August 2018) and compared to GOME-2. The comparison results are shown in **Figure 53**. The results show similar spatial patterns of HCHO column for the two satellites. GOME-2 reports similar HCHO column values as TROPOMI in the same regions.

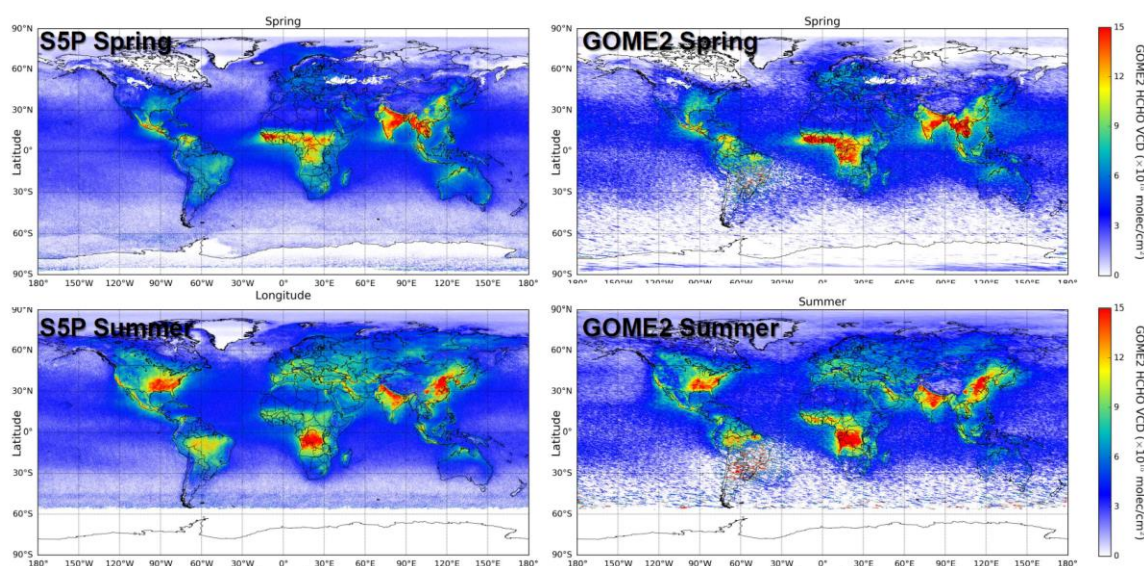


Figure 53: Seasonal average of HCHO vertical column from S5P TROPOMI and MetOp-B GOME 2.

7.3.8 Preliminary validation of UPAS processor upgrade to version 02.01.03

The processor version change to V02.01.03 in July 2020 does not impact the quality of the data. **Figure 54** shows a time series of HCHO differences for five MAX-DOAS stations. The normal behaviour is also observable for example for the two FTIR stations Xianghe and Sodankyla in **Figure 55**.

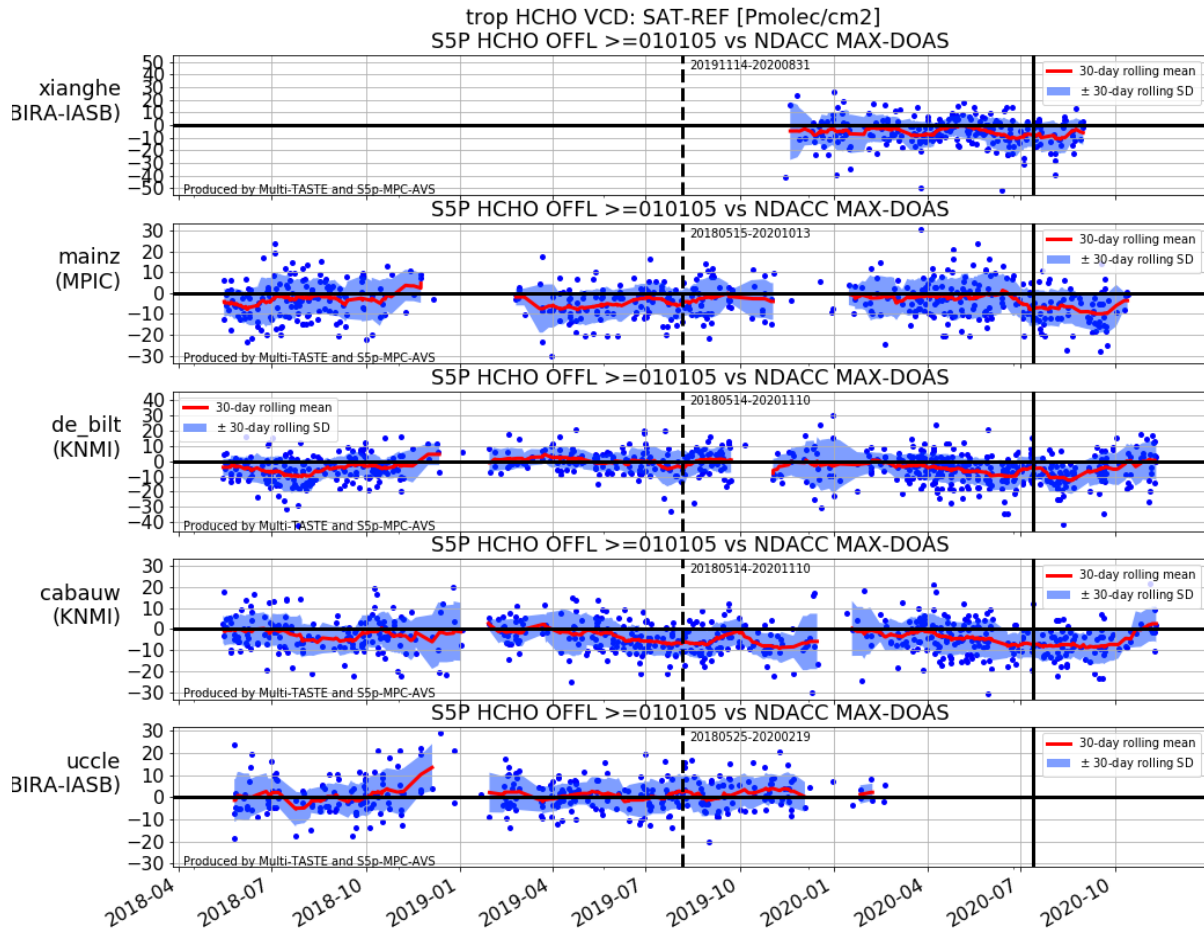


Figure 54: Time series of the difference [Pmolec/cm²] between S5P RPRO+OFFL and MAX-DOAS HCHO column data at five stations. Data was obtained from the VDAF Automated Validation Server on 2020/11/25. Dashed line: pixel size switch at 2019/08/06. Full vertical line: OFFL processor version change on 2020/07/13.

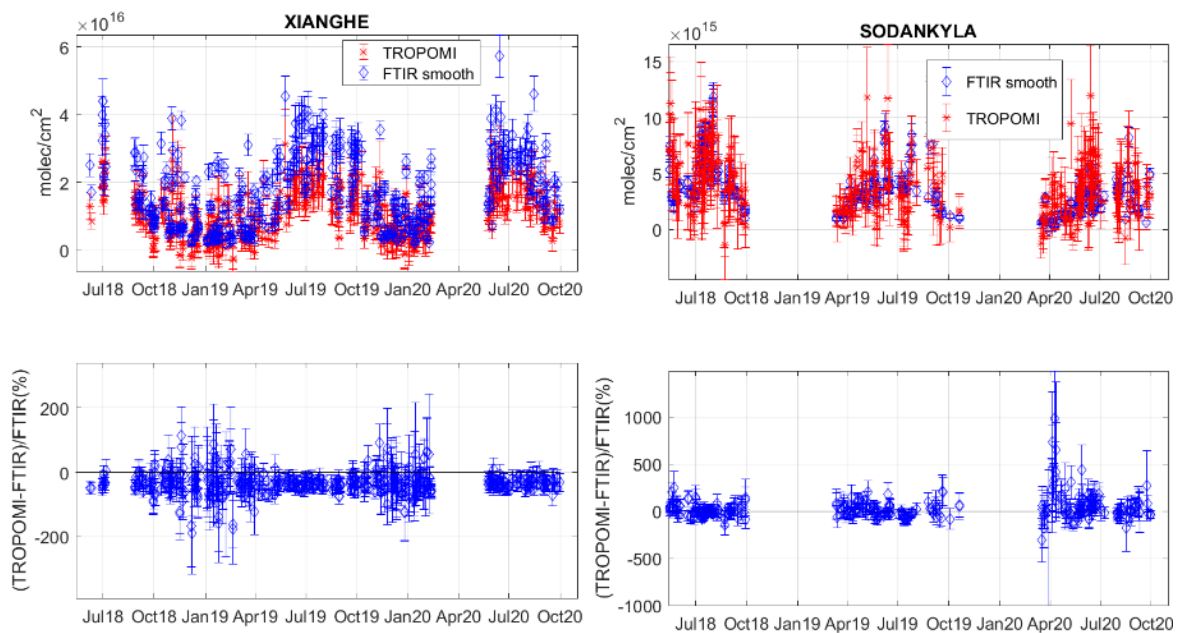


Figure 55: Time series of the HCHO columns [Pmolec/cm²] and difference [%] between S5P and smoothed FTIR HCHO for the stations Xianghe and Sodankyla.

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. L2_HCHO NRTI (V01.01.02 to V01.01.08) and L2_HCHO OFFL (RPRO V01.01.05 and OFFL V01.01.05 to V01.01.08), 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 were determined. The statistical results are summarized in **Table 7**.

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 7**). 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). The daily V02.01.03 difference is -1.2% (± 0.5 Pmolec/cm²) with a Pearson R of 0.96 as shown in **Figure 53**.

7.4.2 Dispersion

Standard deviation and the 1/2 68% interpercentile (1/2 IP68) of the NRTI-OFFL differences are much smaller than that between either NRTI vs MAX-DOAS or OFFL vs MAX-DOAS, indicating a much smaller dispersion between NRTI and OFFL. This is also indicated by the near-unity Pearson R correlation coefficient and slope of NRTI vs OFFL, which are much lower for NRTI vs MAX-DOAS and for OFFL vs MAX-DOAS.

Table 7 – 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 2020-11-25.

Cabauw: 503 common co-locations			
Orbits range from 4839 (2018-09-19) to 15948 (2020-11-10)			
	NRTI vs OFFL	NRTI vs MXD	OFFL vs MXD
Mean(diff)±sem*	-0.09	-4.19±0.37	-4.11±0.37
Median(diff)*	-0.02	-4.76	-4.41
Std(diff)*	1.6	8.2	8.2
1/2 IP68(diff)*	1.2	7.5	7.6
Pearson R	0.98	0.36	0.36
Slope	0.98	0.53	0.52
De Bilt: 482 common co-locations			
Orbits range from 4839 (2018-09-19) to 15948 (2020-11-10)			
	NRTI vs OFFL	NRTI vs MXD	OFFL vs MXD
Mean(diff)±sem*	-0.11	-2.65±0.43	-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	0.24	0.24
Slope	0.96	0.35	0.35

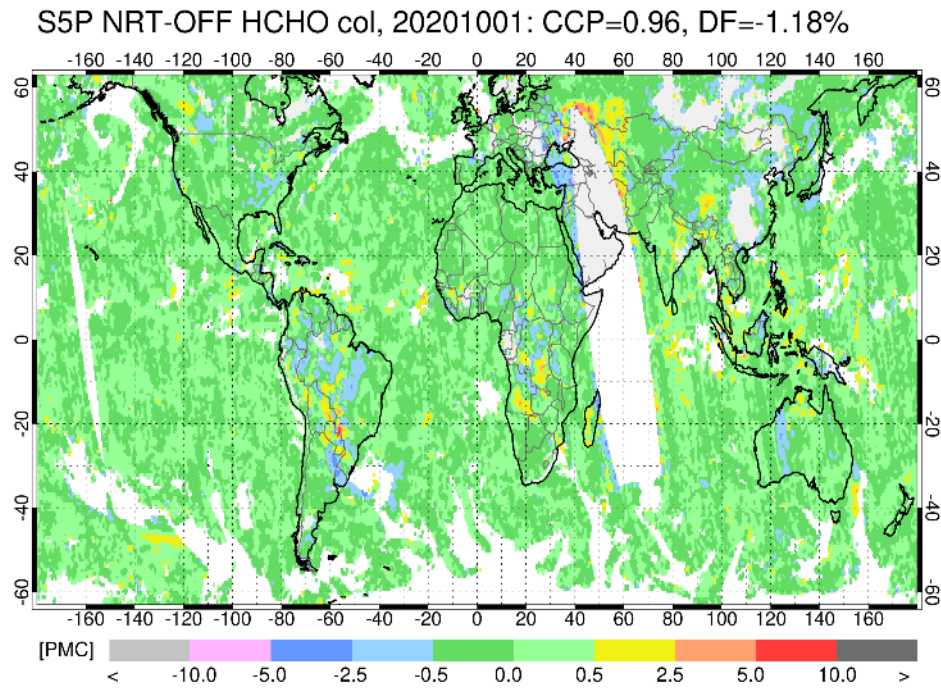


Figure 56: Difference between S5P NRTI and OFFL tropospheric HCHO column data for 2020-10-01 (binned to 0.5°x0.5° resolution). The difference is only calculated for columns above 0.5 Pmolec/cm². Due to the missing NRTI orbit an artefact due to binning is visible.

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. First verification checks of the UPAS processor upgrade to version 02.01.03 released on 16 July 2020 show good consistency of the SO₂ 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 (SO₂ total)

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

Volcanic plumes (SO₂ enhanced)

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

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 SO₂ SCDs (within detection limit) outside volcanic plumes and strongly polluted regions

Outside strongly polluted regions and volcanic plumes, the atmospheric SO₂ concentrations are very low and the corresponding SO₂ 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 SO₂ 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 SO₂, 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, 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, and Basra, Iraq) 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 SO₂ 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 SO₂ columns were not compared directly, because of the strong gradients across volcanic plumes. Instead the derived SO₂ fluxes were compared, for which good agreement was found.

Outside strongly polluted regions and volcanic plumes, the atmospheric SO₂ 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, but the random requirement is often not fulfilled. Here it should be noted that the random requirement is very strict (0.15 – 0.3 DU). For the often very high SO₂ columns in volcanic plumes it is unrealistic that the random requirement can strictly be fulfilled, and it is recommended that the random requirement should be reconsidered;
- from the time series of averaged SO₂ 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 show good consistency of the SO₂ data product before and after the processor switch, with a visible but rather modest effect (seen in **Figure 65** and **Figure 67**).

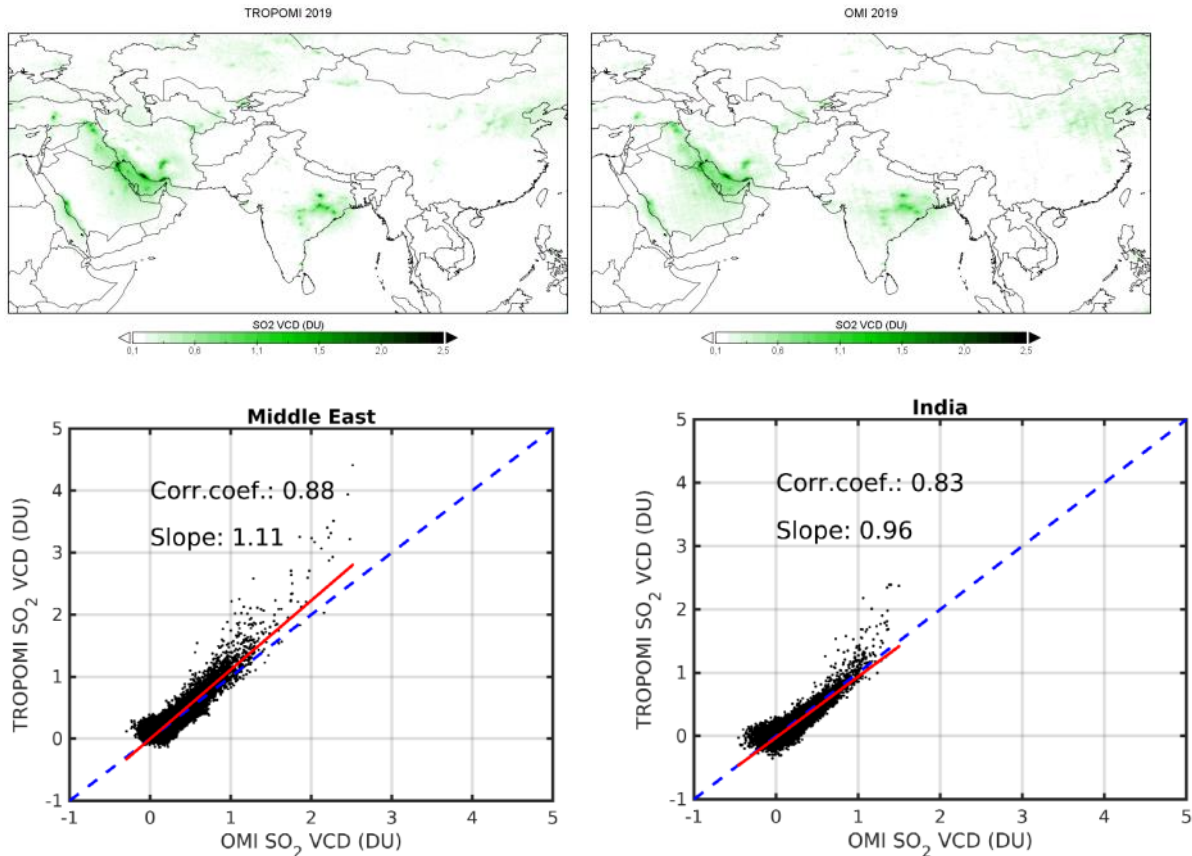


Figure 57: Top: Comparison of the average distribution (01 Jan 2019 – 31 Dec 2019) of the SO₂ VCDs derived from TROPOMI and OMI over regions with strong air pollution. Both data sets show very good agreement. Bottom: Correlation plots TROPOMI versus OMI over the Middle East and India. Note that a fixed AMF of 0.4 was used for both retrievals to exclude the effect of different profile assumptions. Courtesy of Nicolas Theys, BIRA-IASB.

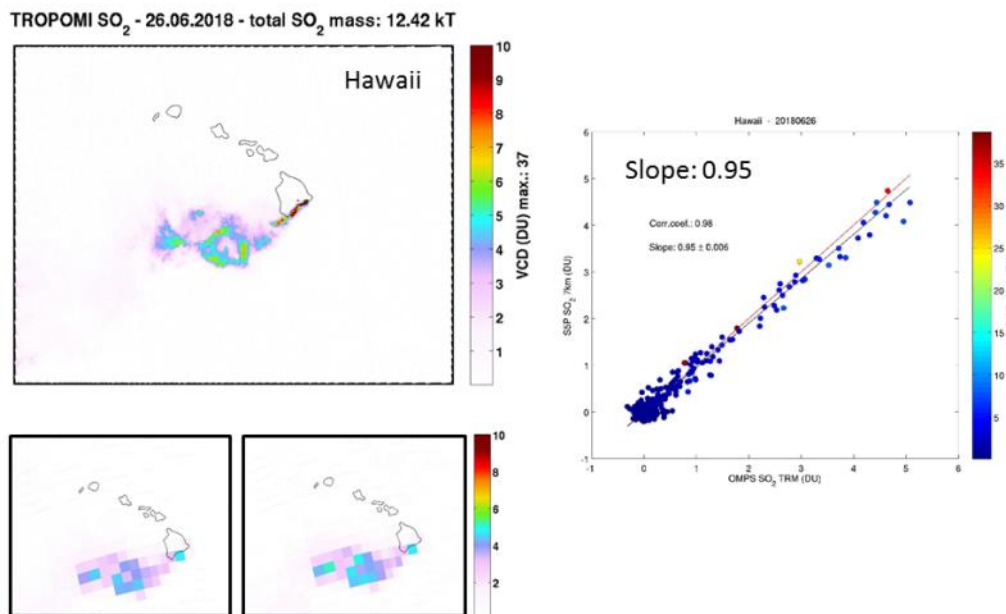


Figure 58: Comparison of TROPOMI and OMPS measurements of the volcanic plume of Kilauea on 26 June 2018. The large figure shows the original TROPOMI data. The two small figures show the spatially degraded TROPOMI data and the OMPS data. The figure right shows the correlation plot of the degraded TROPOMI data versus the collocated OMPS data. Courtesy of C. Li and N. Krotkov, NASA/GSFC.

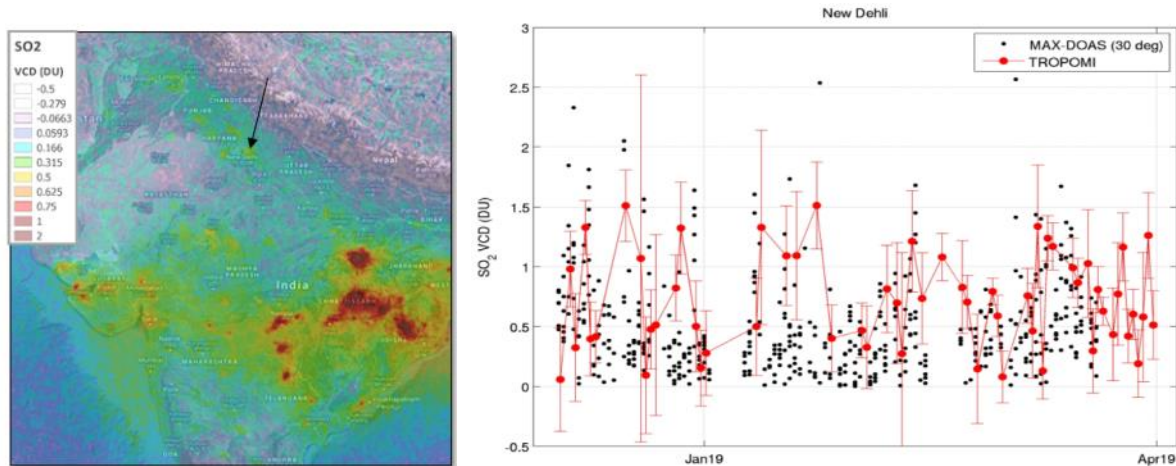


Figure 59: Comparison of TROPOMI SO₂ VCDs to MAX-DOAS measurements at Greater Noida (close to New Delhi, India). The following selection criteria were applied: distance < 15km, CF<0.2, AMF>0.2, MAX-DOAS +/- 1h around S5P overpass. Courtesy of M. Sharma (Sharda University, India) S. Donner, S. Dörner, T. Wagner (MPIC), N. Theys (BIRA-IASB).

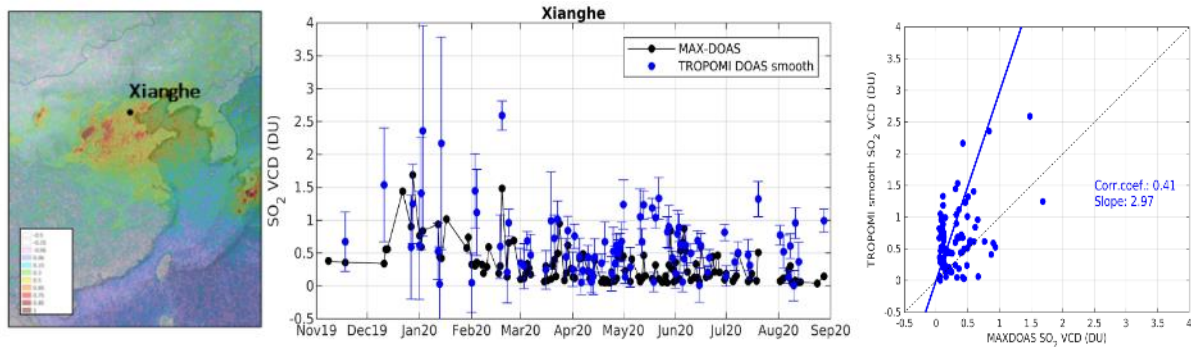


Figure 60: Comparison of TROPOMI SO₂ VCDs to MAX-DOAS measurements (daily means) at Xianghe (China). The following selection criteria were applied: distance < 25km, CF<0.2, AMF>0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB).

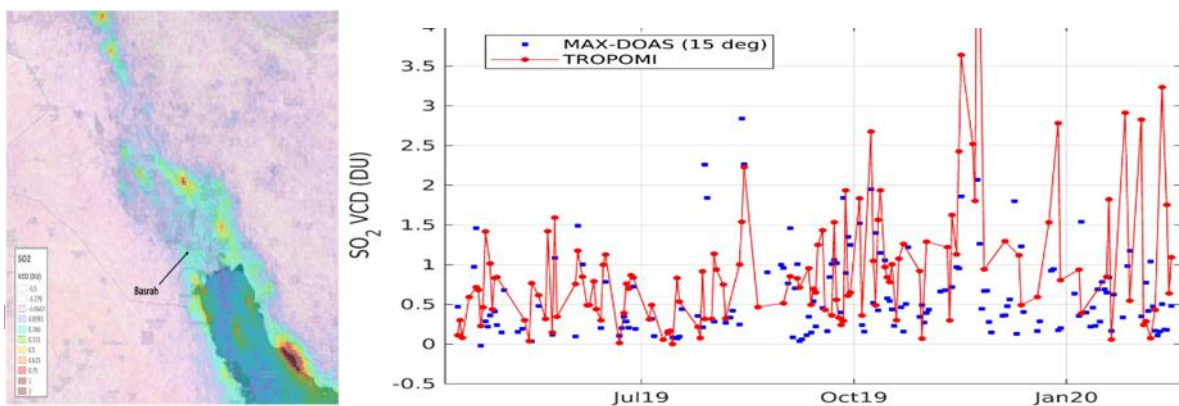


Figure 61: Comparison of TROPOMI SO₂ VCDs to MAX-DOAS measurements (daily means) at Basra (Iraq) from April 2019 to March 2020. The following selection criteria were applied: distance < 25km, CF<0.2, AMF>0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB), data provided by Nayyef Almaliki, Mustafa Aldossary, Ali Almasoudii, Sebastian Donner, Steffen Dörner, Thomas Wagner.

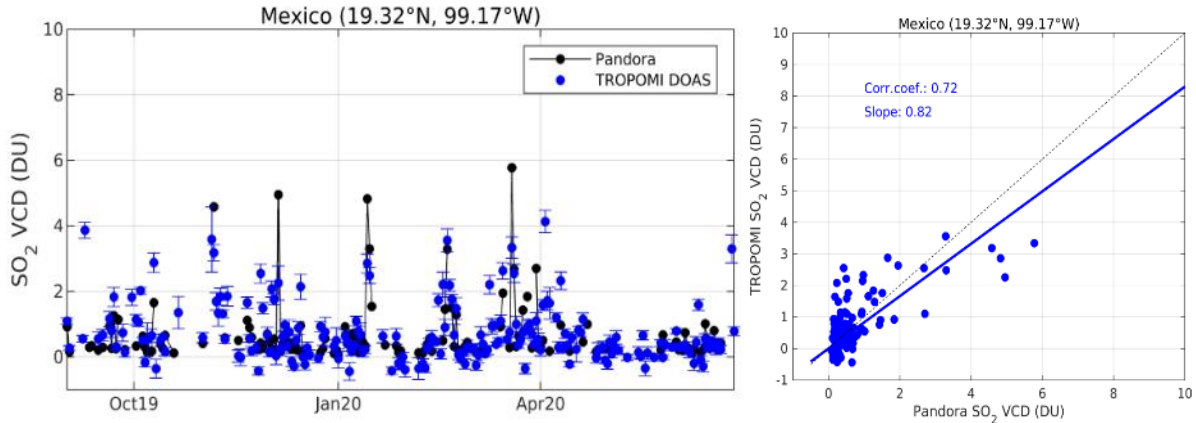


Figure 62: Comparison of TROPOMI SO₂ VCDs to preliminary Pandora measurements (daily means) at Mexico City from September 2019 to August 2020. The following selection criteria were applied: distance < 25km, CF<0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB), Martin Tiefengraber and Alexander Cede (Luftblick). Pandora data are still in demonstration phase.

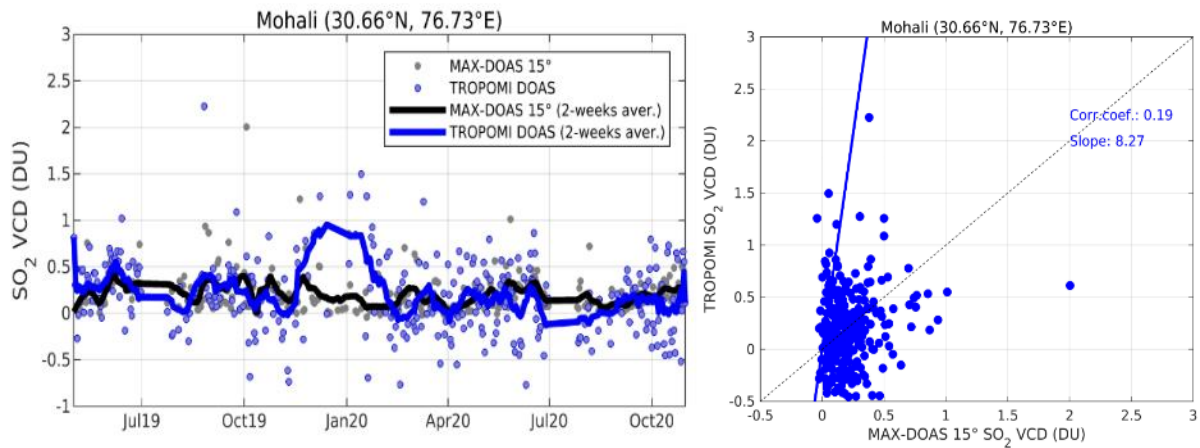


Figure 63: Comparison of TROPOMI SO₂ VCDs to MAX-DOAS measurements (daily means) at Mohali (India) from June 2019 to November 2020. The following selection criteria were applied: distance < 25km, CF<0.2, AMF>0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB), data provided by, Vinod Kumar, Vinayak Sinha, Sebastian Donner, Steffen Dörner, Thomas Wagner.

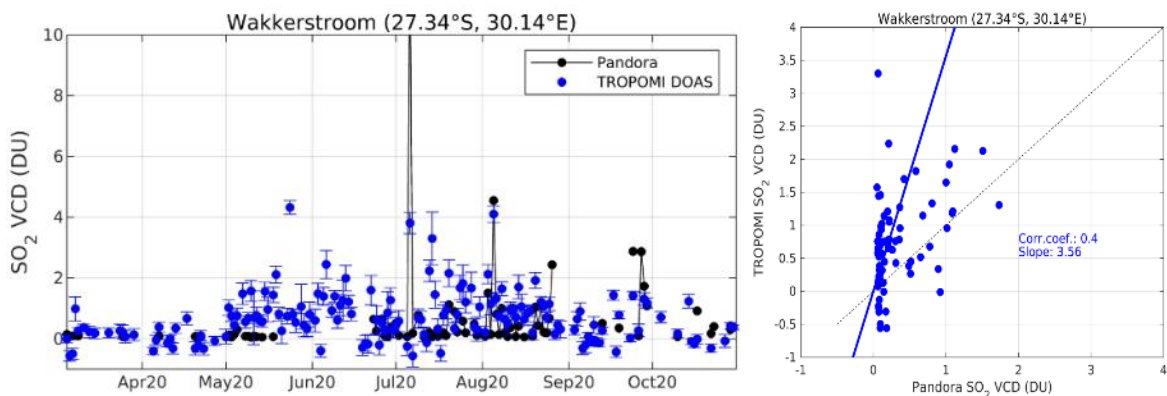


Figure 64: Comparison of TROPOMI SO₂ VCDs to preliminary Pandora measurements (daily means) at Wakkerstroom (South Africa) from March 2020 to November 2020. The following selection criteria were applied: distance < 25km, CF<0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB), Martin 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 SO₂ 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 is slightly above the requirements. However, here it should be noted that the requirements (0.15-0.3 DU) are quite strict and should be reconsidered. 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 SO₂ SCDs it is estimated that the dispersion is within 0.2 DU.

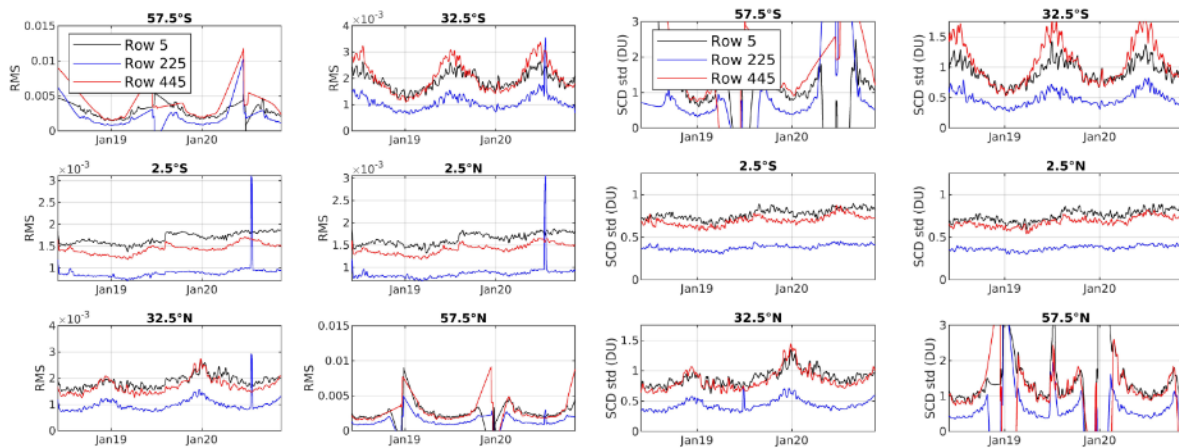


Figure 65: Temporal evolution of the measurement error (left) and the standard deviation (right) for selected 5° latitude bands and 3 detector rows from December 2018 to November 2020. Good qualitative agreement between both quantities is found indicating that the random uncertainty is well characterized 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).

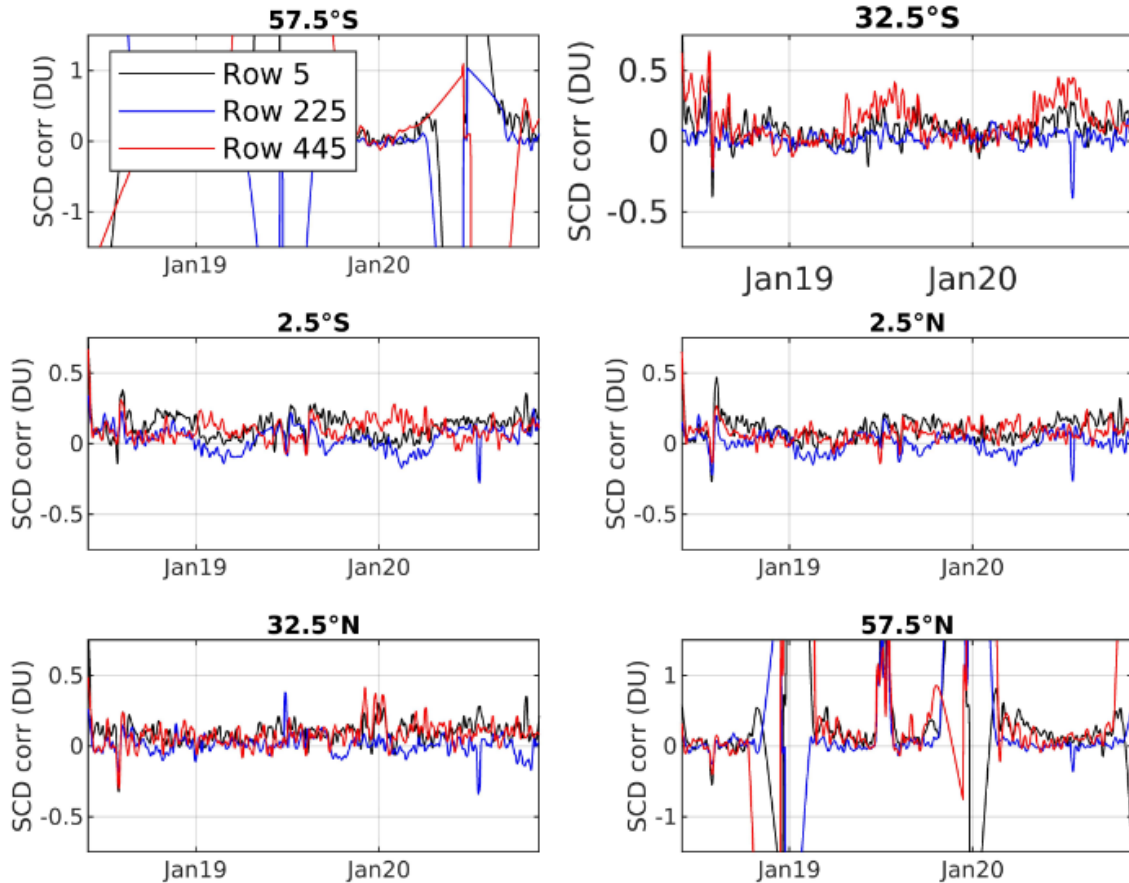


Figure 66: Temporal evolution of the averaged SO₂ SCD for selected 5° latitude bands and 3 detector rows from December 2018 to November 2020. The values are close to zero and show relatively small day to day variations. The larger variations in August 2019 are caused by strong volcanic eruptions. Courtesy 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 SO₂ SCDs (outside periods with strong volcanic eruptions). It is estimated to be 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 SO₂ products are very similar, as illustrated by the comparison of the SO₂ 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.

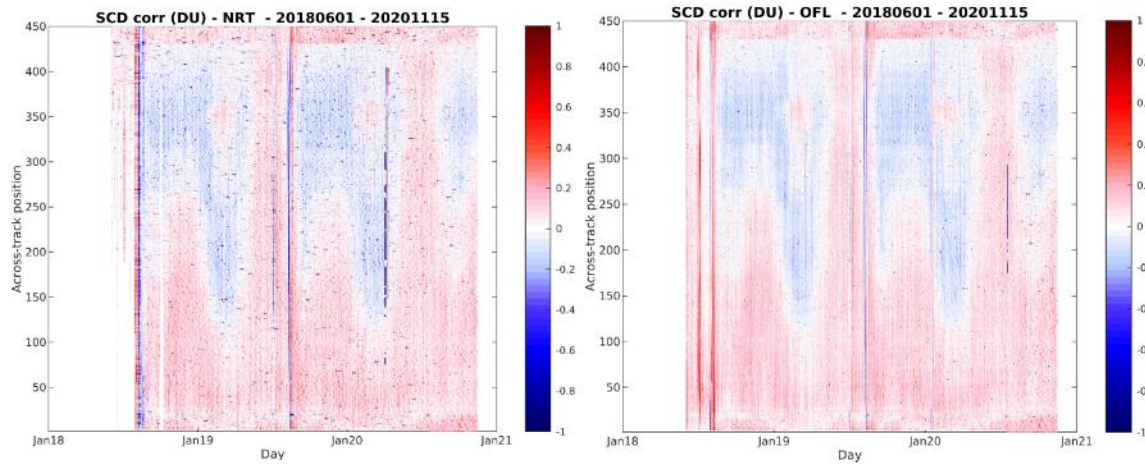


Figure 67: Comparison of the NRT (left) and offline (right) SO₂ data products. Shown are the time series of background corrected SO₂ SCDs for all 450 detector rows from June 2018 to November 2020. Most of the short time features (vertical colored lines) are caused by individual strong volcanic eruptions. 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 68** 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%.

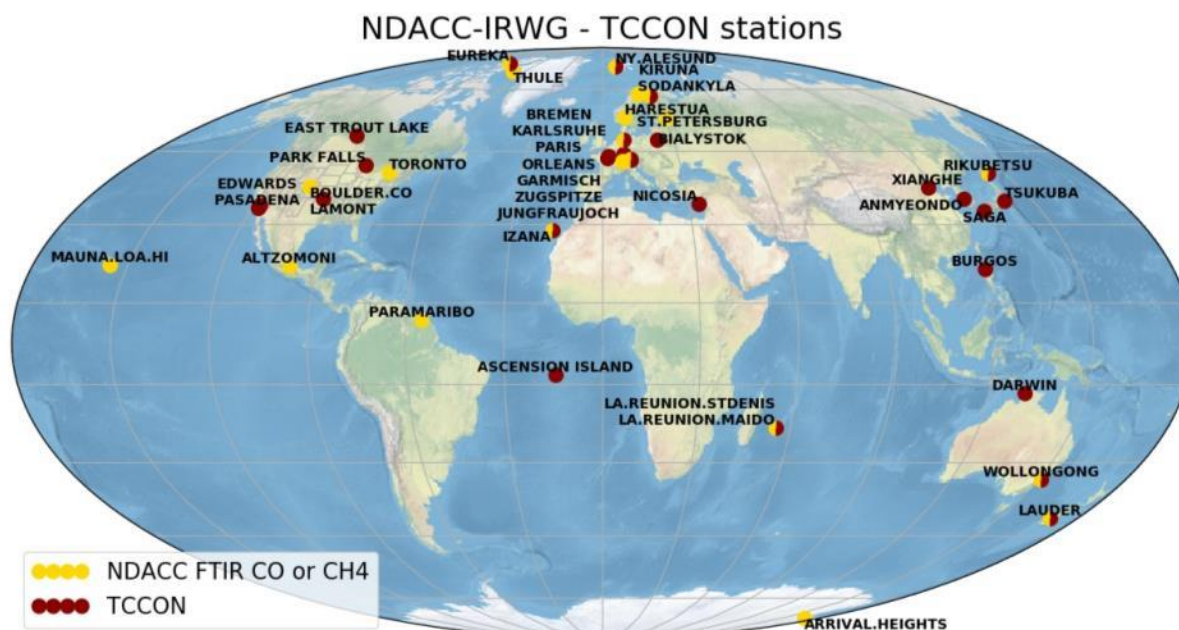


Figure 68: Geographical distribution of NDACC and TCCON FTIR stations measuring atmospheric carbon monoxide column data. Some sites contribute to both networks.

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 NRTI

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 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](http://mpc-vdaf.tropomi.eu) 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 October 2020. 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 9.13% 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 70** and **Figure 71** show the biases over the time period Nov 2017– October 2020 sorted by latitude. It seems that there is no latitudinal dependence of the bias. **Figure 69** does not show any significant degradation in bias with time (note that the longer time period covers different processor versions). **Figure 69** also shows a slight increase of bias during local winter.

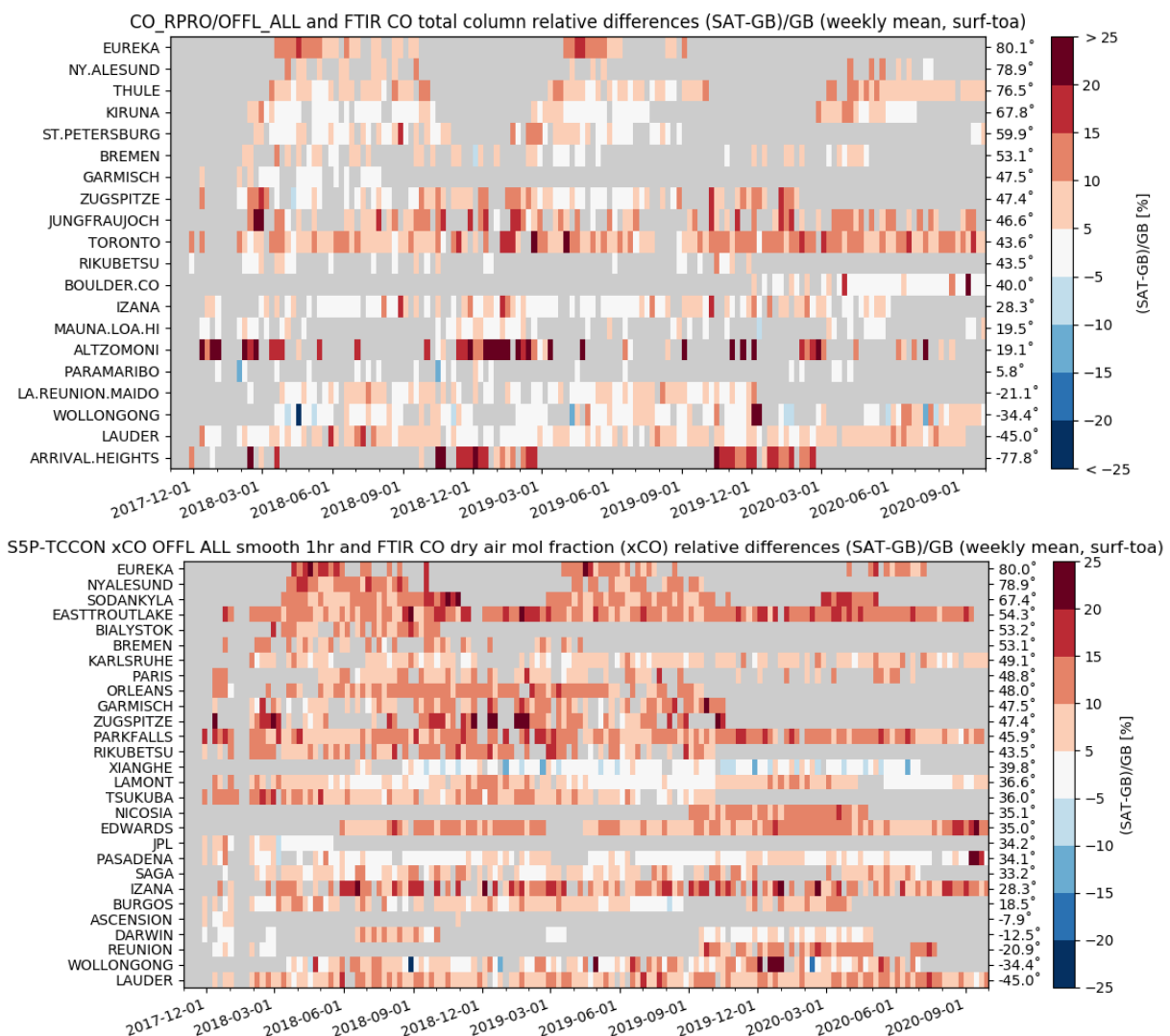


Figure 69: Relative bias between S5P L2_CO OFFL and ground-based CO column data at NDACC (top) and TCCON (bottom) FTIR stations. Over the Nov. 2017 – October 2020 time period the plots do not show a clear meridian dependence or temporal change in the weekly averaged biases.

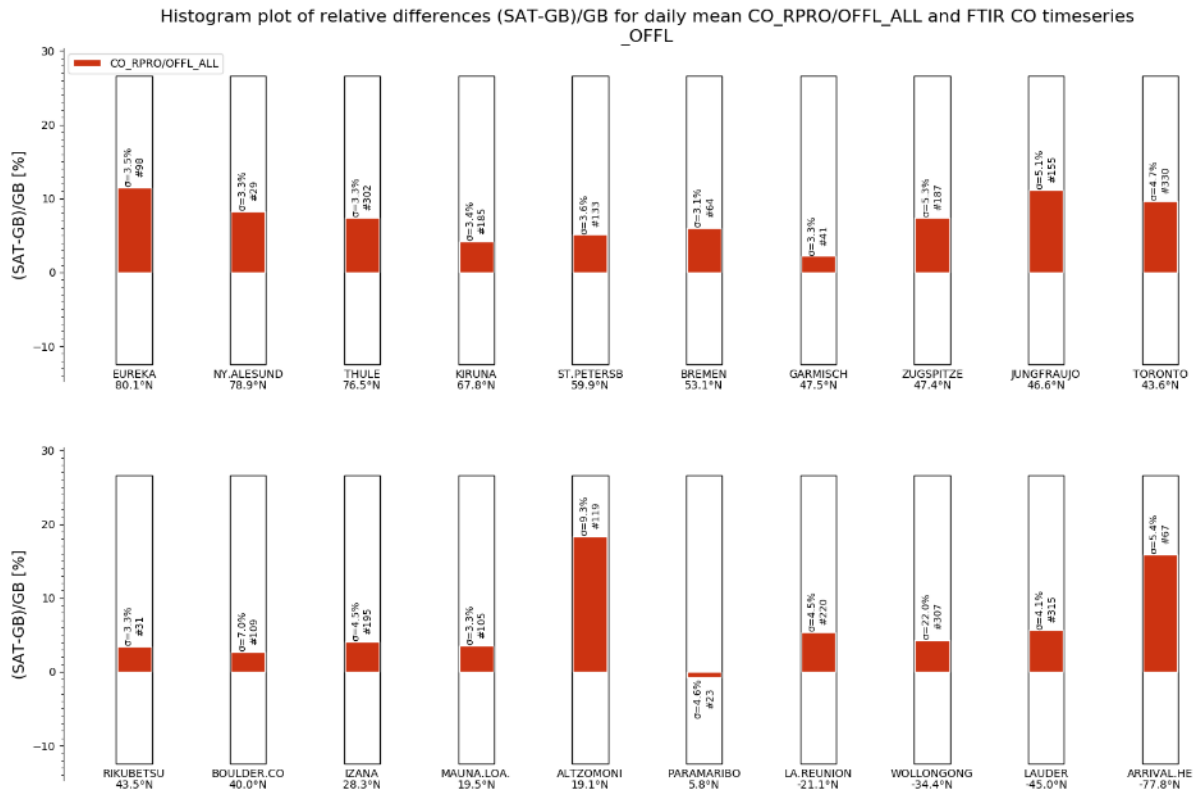


Figure 70: Bar chart of the relative mean difference between S5P L2_CO OFFL and FTIR CO column data at 20 NDACC sites, for all data within the time range from November 2017 till October 2020 showing RPRO/OFFL. The sites are sorted with decreasing latitude. All biases are below 15% except at Altzomoni, 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.

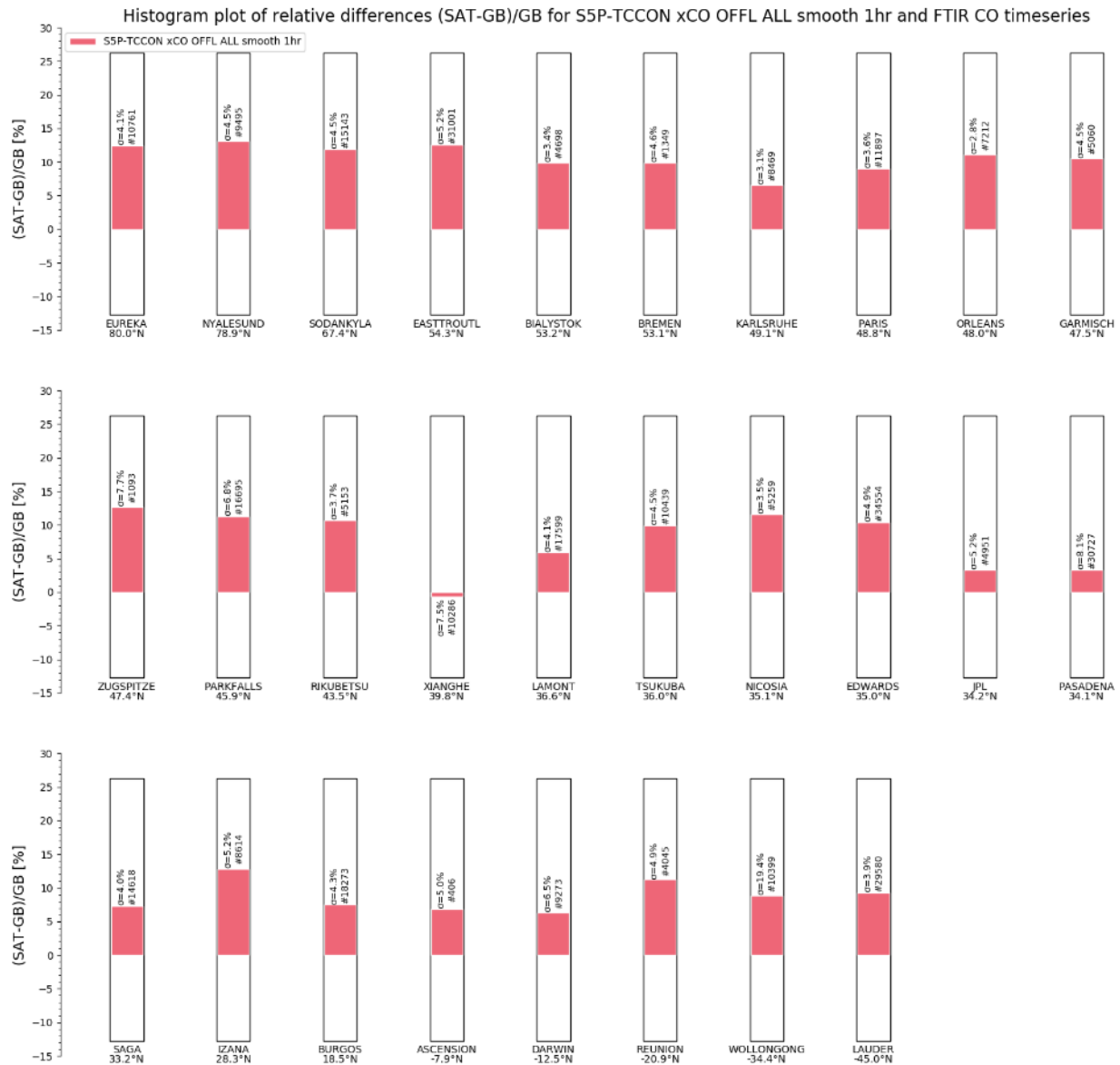


Figure 71: Bar chart of relative mean difference between S5P L2_CO OFFL and FTIR CO column data at 28 TCCON sites for all data within the time range Nov 2017 till October 2020. The sites are sorted with decreasing latitude. The majority of the OFFL biases are below 12% except in the Arctic and mountain 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 70** and **Figure 71**. 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 5% between 10deg and 80deg.

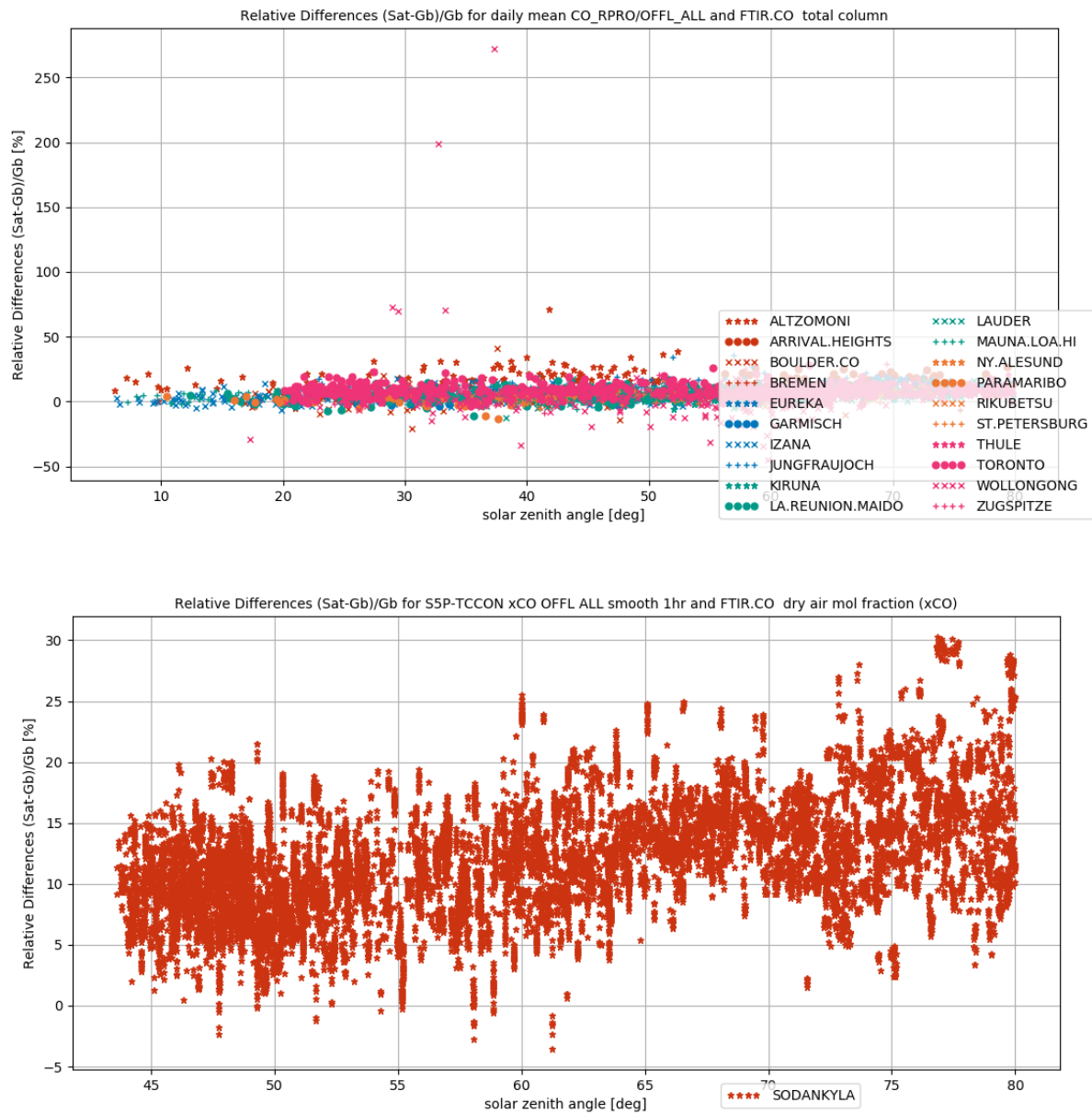


Figure 72: Relative difference (daily mean) between S5P L2_CO RPRO/OFFL and NDACC (top) or TCCON Sodankylä (bottom) carbon monoxide total column as a function of the TROPOMI solar zenith angle, in the 'all' case. The Wollongong time series contains some outlying values which can be attributed to co-location 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 73).

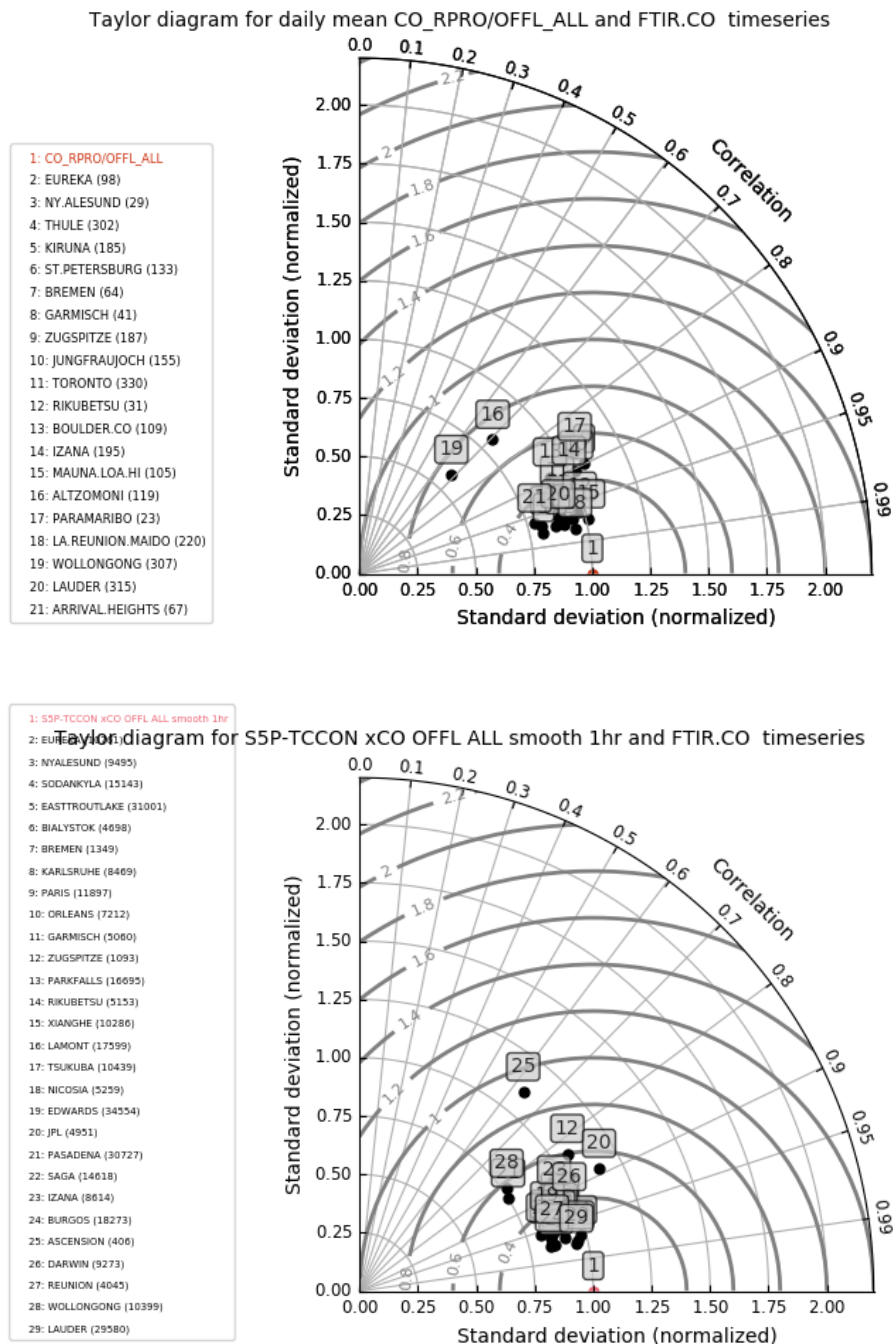


Figure 73: Taylor diagrams for daily mean differences between S5P L2_CO OFFL and ground-based networks CO data: NDACC (top) and TCCON (bottom) for our all case of pixel selection criteria (Boulder and Ascension cover a limited time period), NDACC Wollongong has reduced correlation due to co-location 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 74** 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 75**.

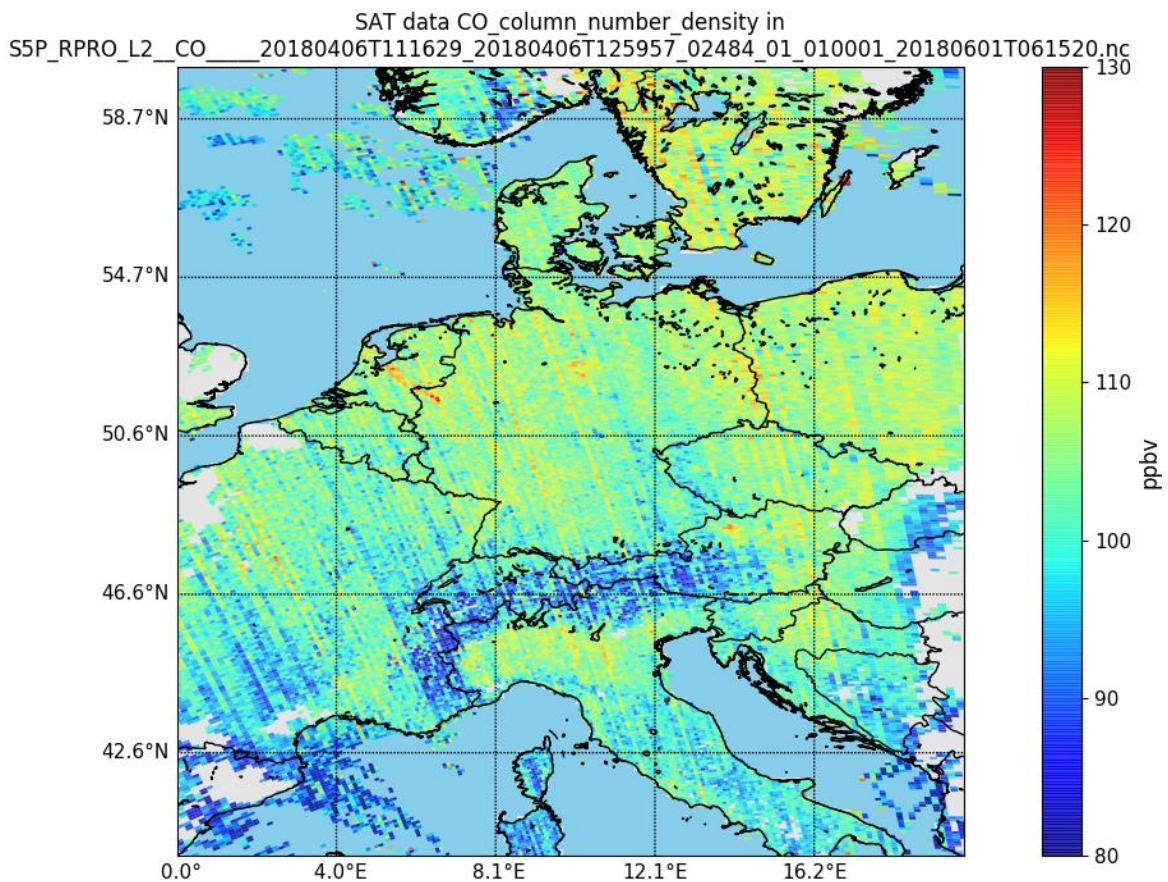


Figure 74: Example of stripe patterns in L2_CO OFFL CO column data along a S5P orbit above Europe.

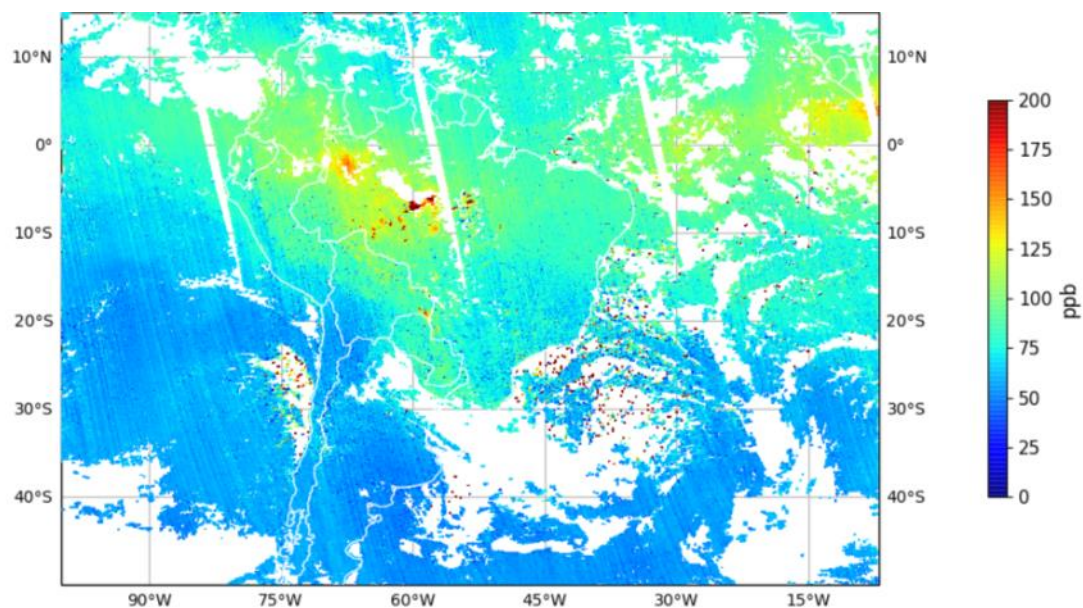


Figure 75: S5p OFFL xCO pixels measured on August 1, 2019, over South America and the Atlantic Ocean. Outlying pixels occurs (including negative values) in the South Atlantic Anomaly.

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

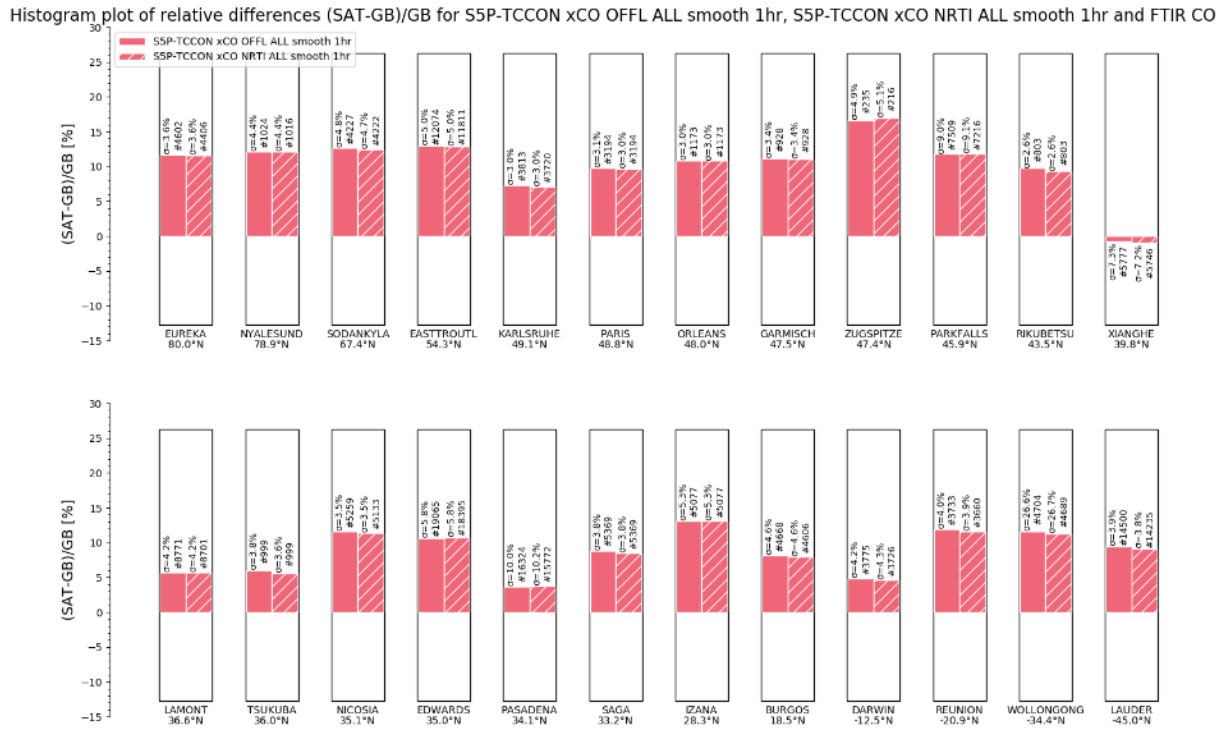


Figure 76: Comparison of relative biases against TCCON CO column data for the S5P L2_CO OFFL and NRTI data versions, from July 3 2019 through October 2020. The quality of both data sets is similar. Over this period the OFFL processor has produced data with a relative mean bias of $9.63\% \pm 5.56\%$ and a correlation coefficient 0.94 with respect to TCCON data; the NRTI processor has produced data with a relative mean bias $9.49\% \pm 5.58\%$ 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://tccodata.org>). **Figure 68** 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](http://mpc-vdaf.tropomi.eu) 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 CH₄ observations more sparse. In the comparison, the apriori in the TCCON and NDACC retrievals have been substituted with the S5P CH₄ apriori (Rodgers 2003). For NDACC data the method described in Rodgers (2003) is followed one step further and the FTIR CH₄ concentration profile (with the S5P prior substituted) is additionally smoothed with the S5P column averaging kernel. The validation procedure for both the NDACC and TCCON based comparisons includes an adaptation of the TROPOMI CH₄ column to the altitude of the ground-based FTIR instrument.

Current conclusions are based on the S5P and reference measurements available at the time of this analysis, which yield comparison pairs from November 2017 through October 2020. Routine validation is done using the Automated Validation Server of the MPC VDAF, the CH₄ validation system operated at BIRA-IASB, and the HARP toolset, and shows an up-to-date comparison.

Since August 6 2019 (orbit 9388) S5P measures with increased spatial resolution from 7km to 5.5km along track. This change in operations did not change the performance of the methane OFFL product.

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.

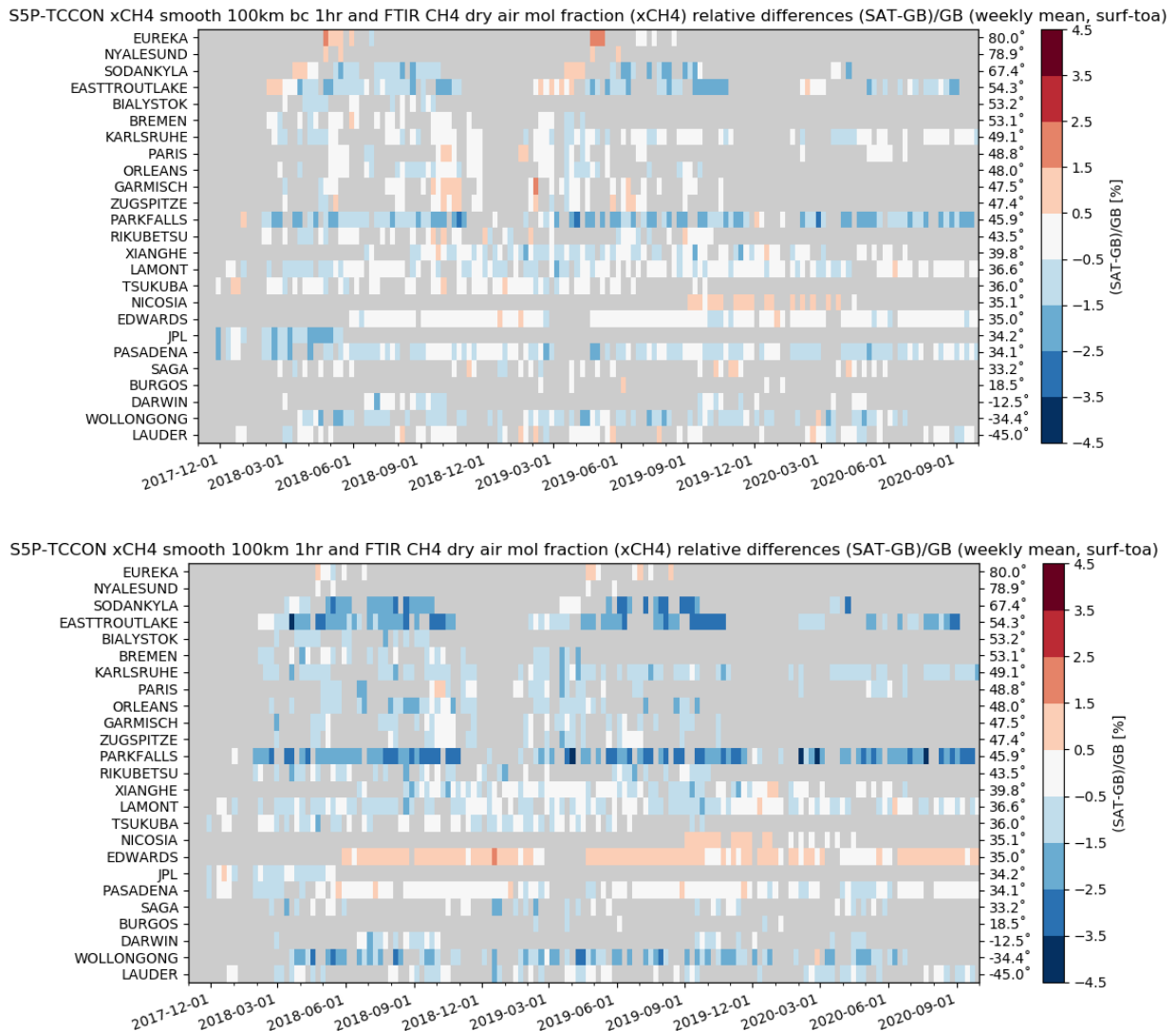


Figure 77: Mosaic plots of relative biases between S5P L2_CH4 RPRO+OFFL and ground-based CH4 TCCON column data for the bias corrected (top) and standard (bottom) methane products. Over the November 2017 – October 2020 time period the plots do not show a clear meridian dependence or temporal change in the weekly averaged biases.

Histogram plot of relative differences (SAT-GB)/GB for S5P-TCCON xCH₄ smooth 100km 1hr, S5P-TCCON xCH₄ smooth 100km bc 1hr and FTIR CH₄ timeseries

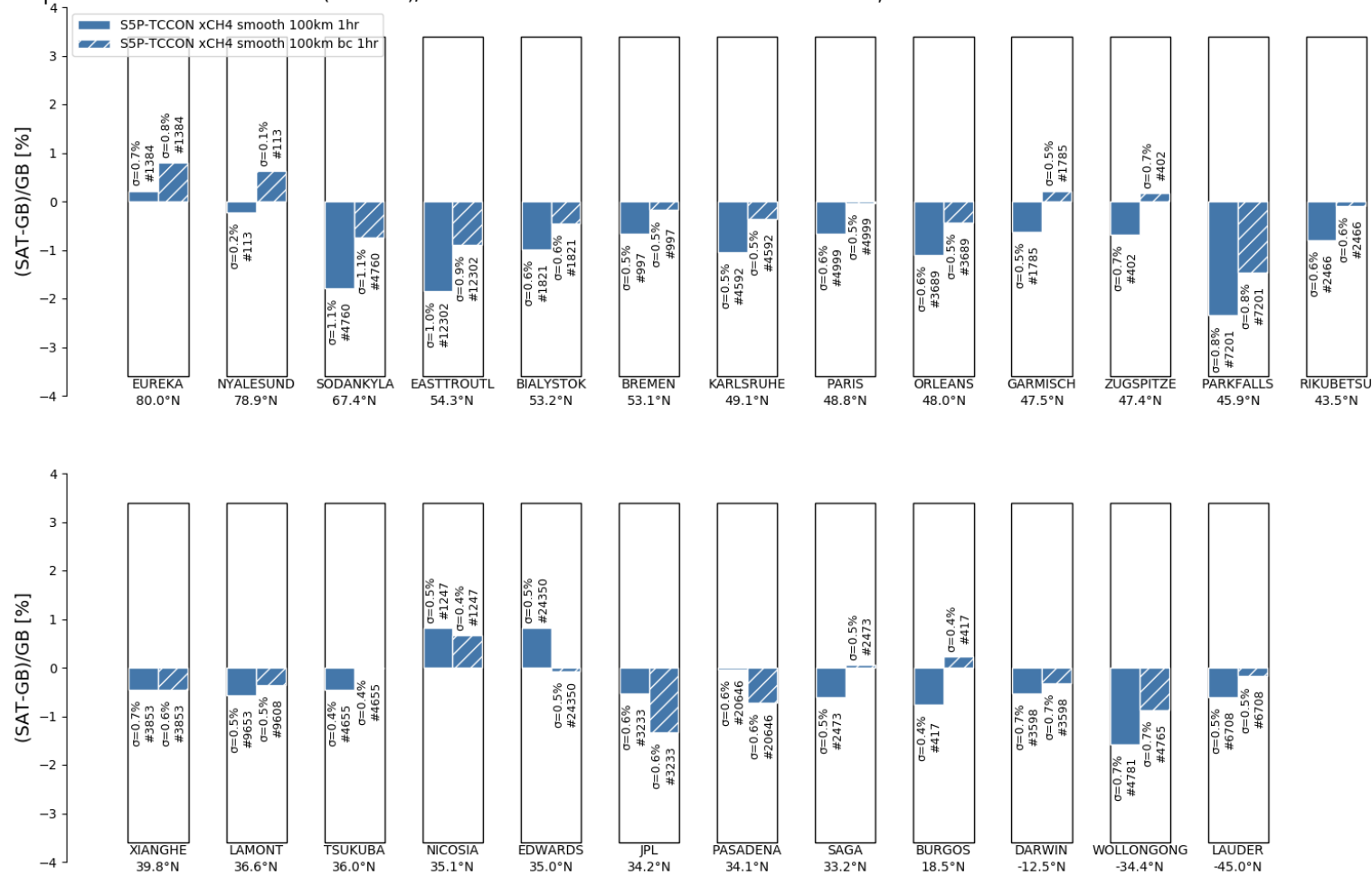


Figure 78: Chart of relative mean difference between S5P L2_CH₄ and FTIR CH₄ column data at 25 TCCON sites within the time range November 2017 till October 2020. The sites are sorted with decreasing latitude. The relative mean difference of the standard xCH₄ product slightly exceeds the mission requirements (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.

site	S5P-NDACC xCH4 smooth 100km 3hr					S5P-NDACC xCH4 smooth 100km bc 3hr					lat
	#	rel. NDACC std	correlation	rel diff bias(%)	rel diff std(%)	#	rel. NDACC std	correlation	rel diff bias(%)	rel diff std(%)	
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	0.89	3.57	1.07	4	2.7	0.97	4.54	1.02	78.9
THULE	67	1	0.74	3.39	0.8	67	0.9	0.77	4.45	0.78	76.5
KIRUNA	110	0.9	0.09	-1.09	1.44	110	0.9	0.09	-0.11	1.49	67.8
SODANKYLA	116	1.1	0.3	-0.96	1.16	116	1	0.34	0.08	1.14	67.4
ST.PETERSBURG	96	1.2	0.39	-0.7	0.94	96	1.2	0.5	0.26	0.85	59.9
BREMEN	43	2.2	0.49	0.79	1.49	43	2.3	0.52	1.33	1.47	53.1
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	-0.02	0.62	60	1	0.69	0.82	0.61	47.4
JUNGFRAUJOCH	58	1	0.65	-0.77	0.75	58	1	0.67	-0.11	0.74	46.6
TORONTO	109	2.8	-0.02	0.95	2.93	109	2.6	-0.06	1.53	3.01	43.6
RIKUBETSU	23	2.6	0.08	1.15	2.29	23	2.5	-0.01	1.91	2.4	43.5
BOULDER.CO	76	1.2	0.45	2.15	0.79	76	1.1	0.35	1.77	0.87	40
ALTZOMONI	58	1	0.17	2.54	0.81	58	1	0.38	2.37	0.73	19.1
WOLLONGONG	32	1.4	0.57	-0.59	0.83	32	1.4	0.58	0.15	0.83	-34.4
LAUDER	70	1.3	0.7	-0.26	0.83	70	1.4	0.64	0.17	0.89	-45
ARRIVAL.HEIGHTS	21	0.8	0.62	1.76	0.89	21	0.8	0.6	2.68	0.89	-77.8
		1.5	0.47	0.68	1.14	--	1.4	0.49	1.36	1.15	

Table 8 – 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.).

10.3.6 Short term variability

For all the NDACC and TCCON stations, short scale temporal variations in the CH₄ column as captured by ground-based instruments are reproduced very similarly by S5P L2_CH4 OFFL. The individual Pearson correlation coefficients are on average 0.67, see **Figure 80**. 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.

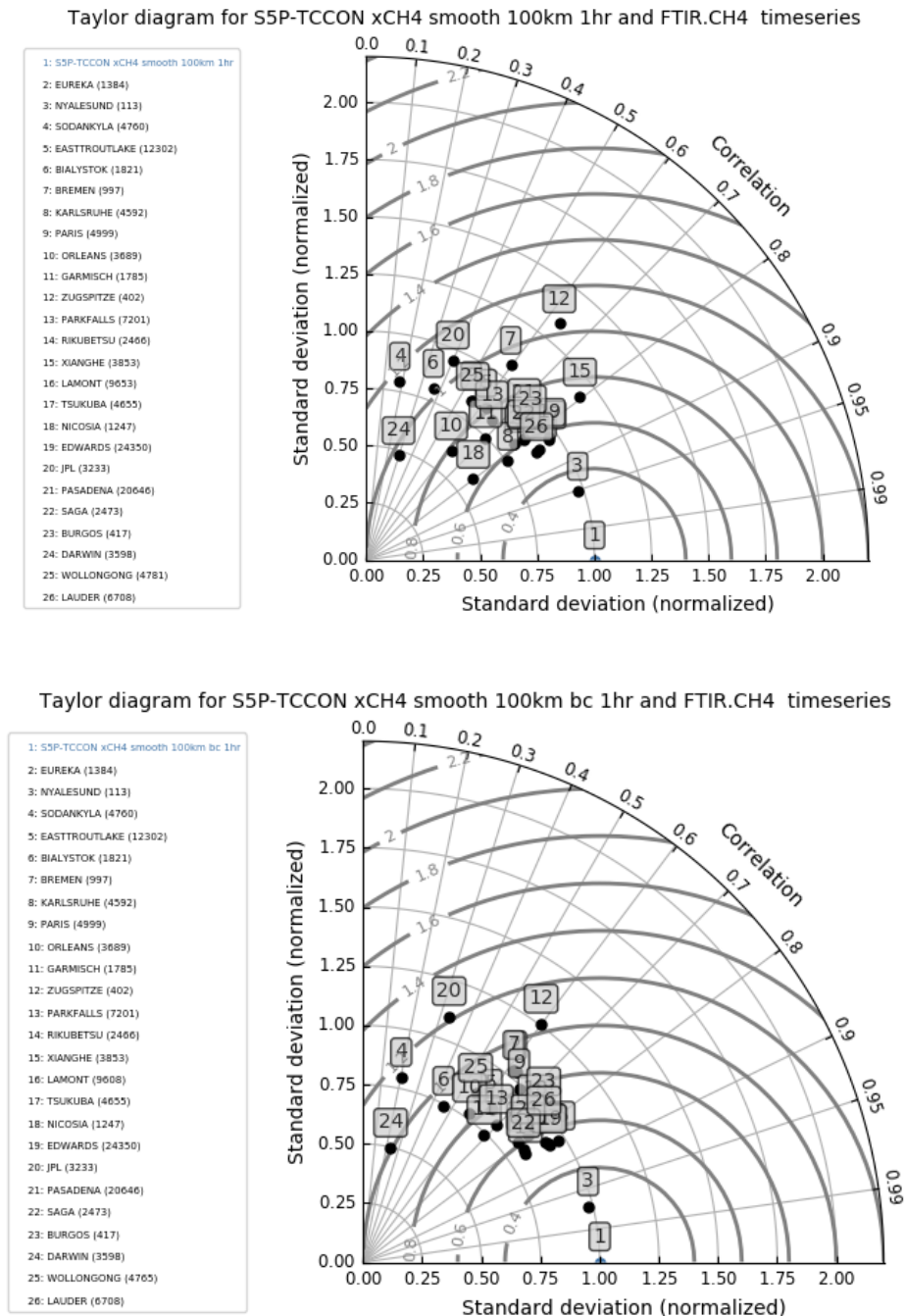


Figure 80: Taylor diagrams for differences between S5P L2_CH4 RPRO/OFFL and TCCON methane column data: standard (top) and bias corrected (bottom) S5P methane columns. At almost all sites the variability of the S5P column data is higher compared to the variability in the TCCON data.

10.3.7 Geographical patterns

Single TROPOMI overpasses show stripes of erroneous CH₄ values in the flight direction (see **Figure 81** left). For orbits before orbit 7644 (April 5 2019) not all pixels above inland water are filtered with the qa_value flag, see **Figure 81** (right, above Caspian Sea).

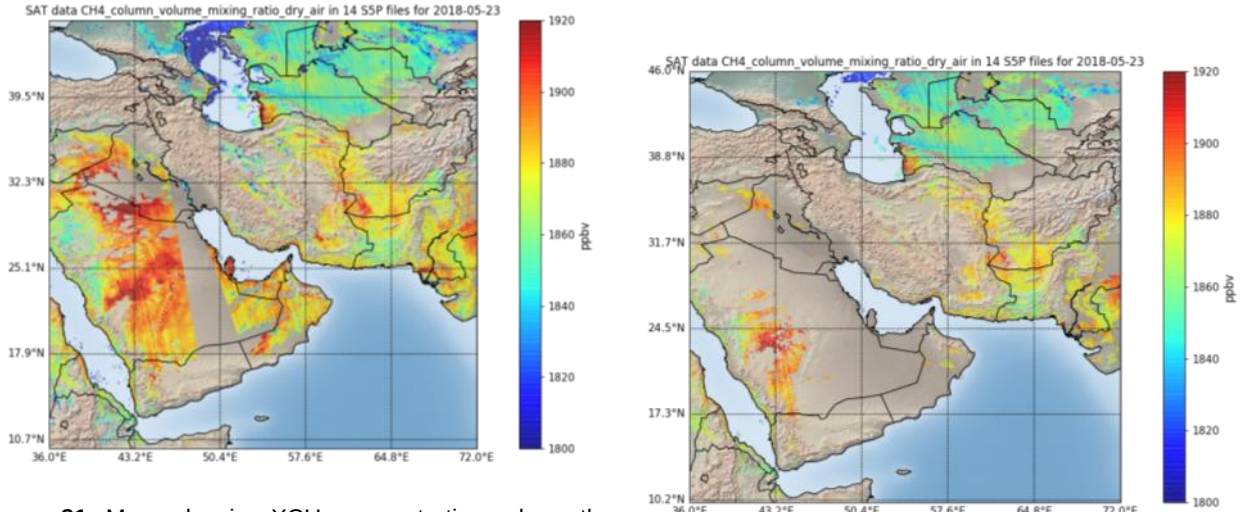
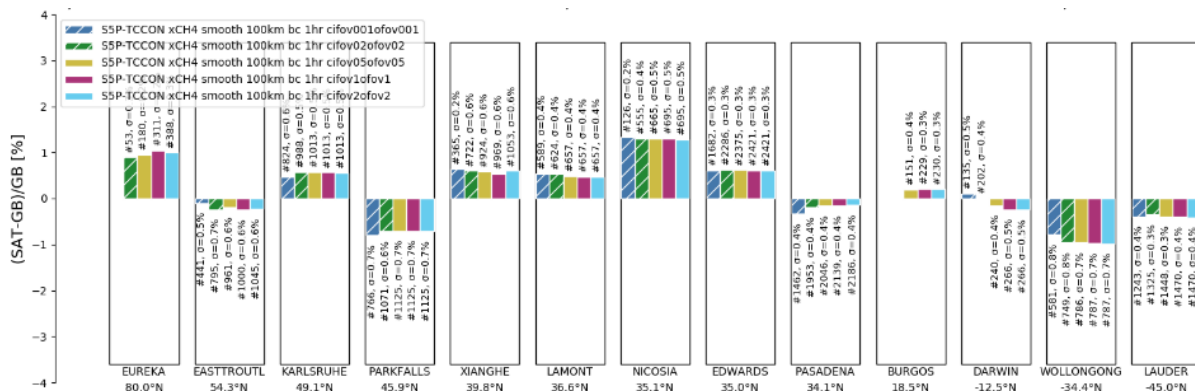


Figure 81: Maps showing XCH₄ concentrations above the Middle East measured by S5P on May 23 2018. The left panel shows all available ground pixels, the right panel shows only pixels with qa_value>0.5. The left panel shows the presence of stripes in the flight direction and the right panel shows the presence of filtered pixels above inland water (Caspian Sea).

10.3.8 Change in Cloud Mask

Since March 11 2020 the cloud data used for the CH₄ data quality filtering has changed (cf PDGS release V2.6.0). The ECM Cloud Mask is used instead of VICMO since orbit 12432. The qa_value threshold of 0.5 removes significantly more pixels since March 11. This has resulted in a reduced number of co-located days and pixels that are found for the NDACC and TCCON stations. The XCH₄ data earlier than March 11 2020 uses a cloud fraction (CF) prefilter 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 82**). A CF threshold of 0.001 was proposed by the L2_CH4 team and TCCON validation results supports this change. The cloud fraction definition will be modified in the next processor update.



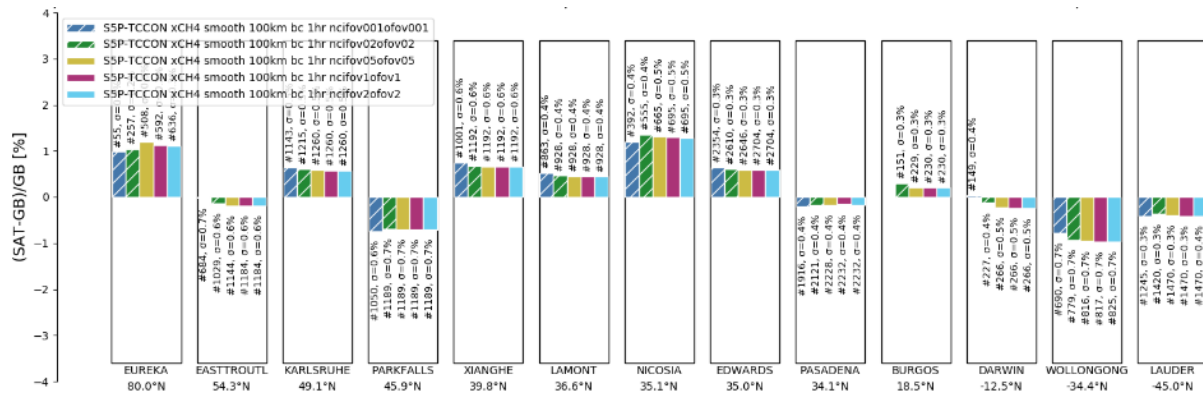


Figure 82: charts of relative mean difference between S5P L2_CH4 and FTIR CH₄ column data at 13 TCCON sites within the time range 11 March 2020 till 19 July 2020. The sites are sorted with decreasing latitude. Top panel plot is for the old CF definition and the bottom panel plot is for the new CF definition. Different colors 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 83**).

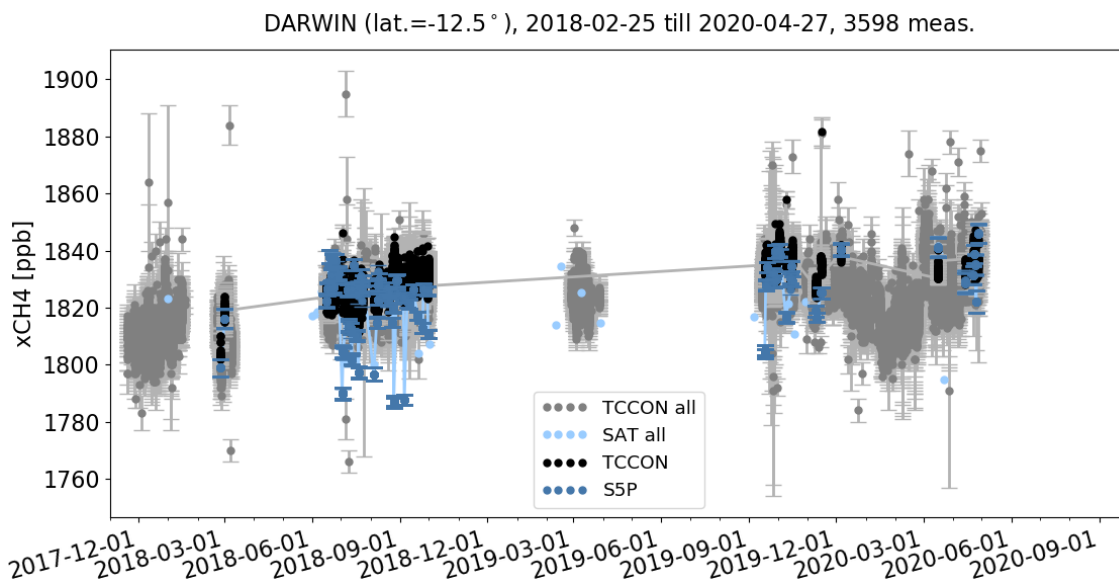


Figure 83: S5P L2_CH4 XCH₄ time series over Darwin where low values of XCH₄ are observed for several days.

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 Cloud Fraction (CF), the Cloud Height (CH), and the Cloud Optical Thickness (COT). Validation results are discussed with respect to the product quality targets outlined in **Table 3**. Before the UPAS processor switch to version 02.01.03 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 11.5 demonstrates evidence that NRTI and OFFL data do not differ significantly and that their respective validations yield similar conclusions before the processor switch. Subsection 11.4 reports on an initial validation of the UPAS processor upgrade to version 02.01.03.

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 9**) 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 9 – 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 O₂-O₂

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

¹ The VIIRS cloud datasets were obtained from a pre-production code run specifically for limited S5P team analysis. The VIIRS cloud algorithm is based on the MODIS Collection 6 algorithms [<https://modis-atmosphere.gsfc.nasa.gov/documentation/collection-6>; Platnick et al. (2017)]. The CLDPROP data were released in Feb. 2019 and are described here: <https://modis-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 O₂ A band. As opposed to CLOUD CRB, where cloud albedo is retrieved, in FRESCO-S, the cloud albedo is assumed to be 0.8 or the TOA reflectance at 758 nm if the reflectance is larger than 0.8. We note that at small cloud fractions, the surface albedo is adapted to prevent negative cloud fractions.

Finally, CLOUD CAL cloud top height is compared with CLOUD CRB cloud height at the CLOUDNET sites.

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: $CF_rescaled = CF \cdot CA / 0.8$.

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>75°), inhomogeneous scene warnings, high residual of the fitting process etc.

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

1. instrumental feature: spatial mis-registration between TROPOMI bands 3-4 (OCRA, UV trace gas fitting window) and band 6 (ROCINN fitting window),
2. insensitivity to very thin clouds,
3. treatment of multi-layer clouds,
4. treatment of ice clouds,
5. snow/ice conditions,
6. unknown straylight impact in the NIR,
7. saturation (note that the L1B flagging works well, only blooming isn't flagged correctly yet),
8. some ground pixels contain cloud-height values close to the a-priori. This behavior is related to the current setting of the inversion algorithm. This bug is resolved from version 01.01.06 onwards.
9. Version 01.01.06 had an inconsistency in cloud parameters; for pixels with a priori cloud fraction below 0.05, the cloud height and other properties were set to fill values which caused data gaps in the ozone product. This problem is corrected in 01.01.07 by setting cloud fractions below 0.05 to 0.0 in the retrieval. The original a priori cloud fraction is maintained in the variable `cloud_fraction_apriori`.

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>.

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 the *S5P First Public Release Validation Workshop* (ESA/ESRIN, June 25-26, 2018), the *3rd S5PVT workshop* (ESA/ESRIN, November 11-14, 2019) and the S5PVT meeting at October 2020. 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, while 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 01.01.07-02.01.03 data. This covers the time period from 2018-04-30 till November 2020. 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

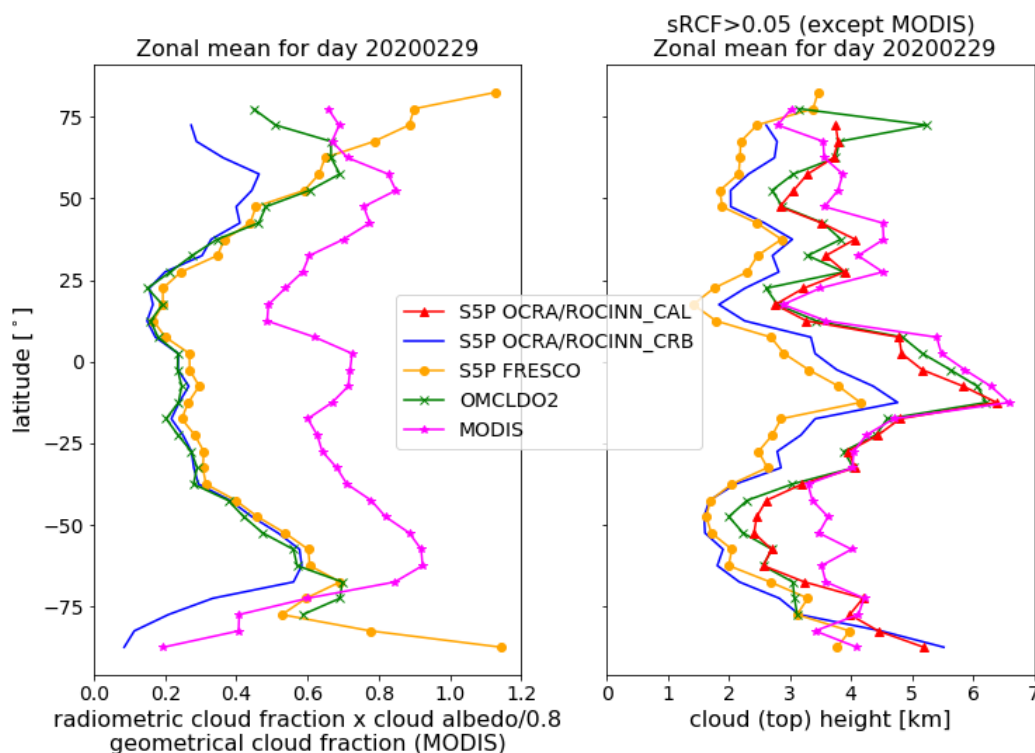


Figure 84. Zonal mean plots (5 deg latitude grid) for 2020-02-29 of radiometric cloud fraction scaled to a cloud albedo of 0.8 of S5P CLOUD CRB, S5P FRESCO and OMI OMCLDO2, and the geometrical cloud fraction of MODIS. Right: cloud (top) height of S5P CLOUD CAL, S5P CLOUD CRB, S5P FRESCO, OMI OMCLDO2 and MODIS. Note that in the right panel, the pixels with scaled radiometric cloud fraction lower than 0.05 are removed (does not apply to MODIS).

The latitudinal variation of zonal means of radiometric cloud fraction, rescaled to a cloud albedo of 0.8 is compared for several satellite products (**Figure 84**, left panel) for 2020-02-29. Comparison with alternative algorithm S5P FRESCO. The scaled cloud fraction of S5P CLOUD CRB and of OMI OMCLDO2 match closely between approximately -60 deg and +40 deg latitude. The geometrical cloud fraction of MODIS is also provided. As expected, it is shifted compared to the S5P CLOUD CRB scaled cloud fraction, but has a similar latitudinal variation.

Comparison with alternative algorithm S5P FRESCO

Radiometric cloud fraction from L2 CLOUD OFFL CRB and from the support product L2 FRESCO were first rescaled to cloud albedo of 0.8 and then compared. Boxplots of the difference CRB-FRESCO and the normed relative difference $(\text{CRB-FRESCO})/[0.5 \cdot (\text{CRB} + \text{FRESCO})]$.

Figure 85 presents boxplots of the difference (top) and normed relative difference (bottom) between rescaled cloud fractions of CLOUD CRB and FRESCO at the CLOUDNET sites. Overall, mean and median difference are negative (CRB lower than FRESCO) but 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 Mado (Reunion) a stronger deviation is seen.

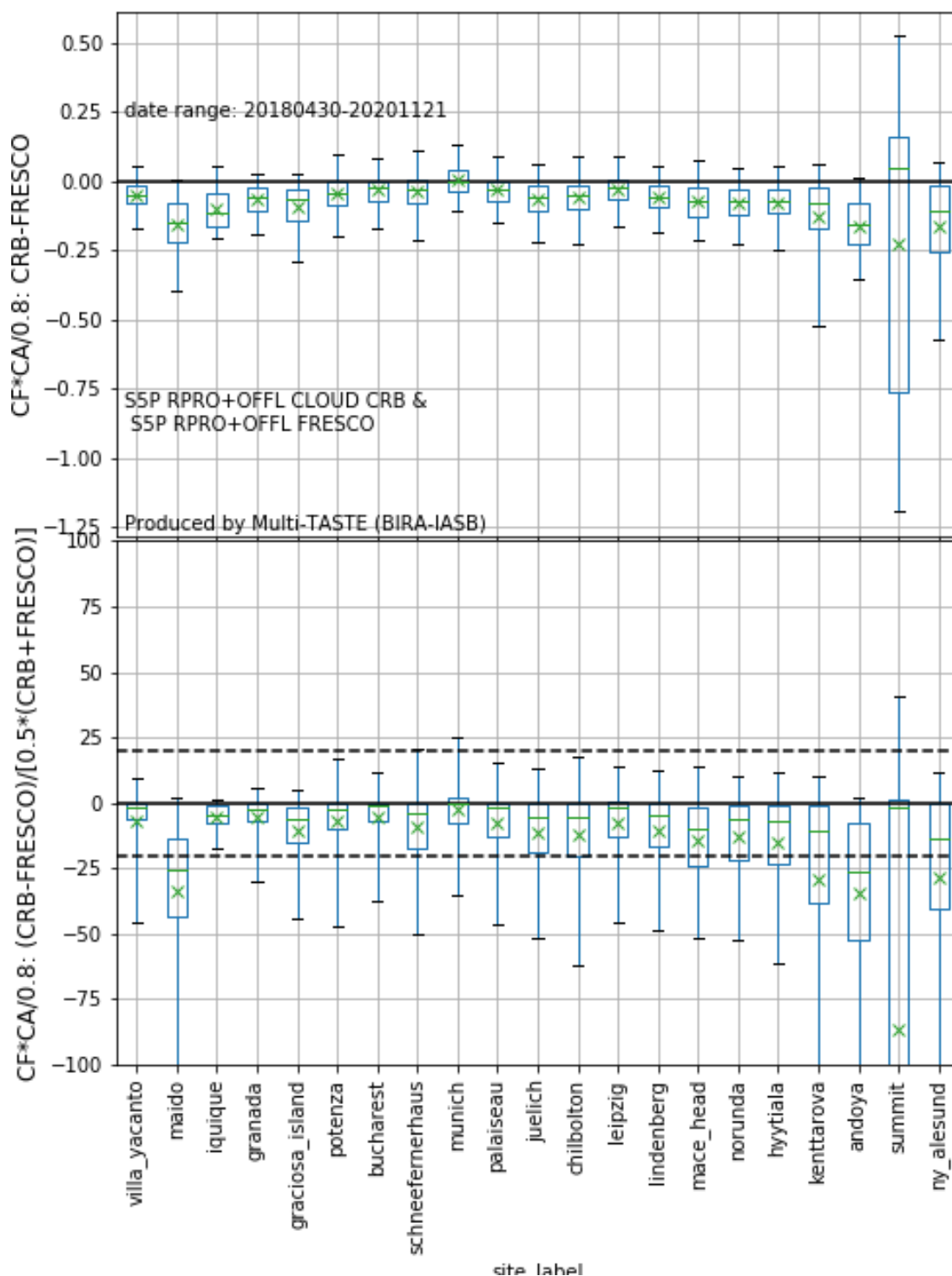


Figure 85. Top. Box plots of S5p CLOUD CRB cloud fraction minus S5p FRESCO cloud fraction, after both have been rescaled to a cloud albedo of 0.8. Bottom: same but for the normed relative difference.

In the zonal mean plot (Figure 84) S5P FRESCO scaled radiometric cloud fraction matches well between approximately -60 deg and +40 deg latitude, although slightly shifted towards a higher value. At the more extreme latitudes S5P CLOUD CRB and S5P FRESCO diverge, likely due to a different treatment of snow-ice cover.

11.3.3.2 Dispersion

Figure 86 presents two dispersion measures of the CRB minus FRESCO rescaled cloud fractions: standard deviation (left) and the more robust $\frac{1}{2}$ IP68. While the standard deviation is close to 0.1 at most sites, the $\frac{1}{2}$ IP68 is lower and only slightly exceeding the dispersion requirement of 0.05. The high dispersion at the high-latitude sites Kenttarova, Andoya, Summit and Ny-Alesund (**Figure 85**, **Figure 86**) 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 Mado (Reunion Island).

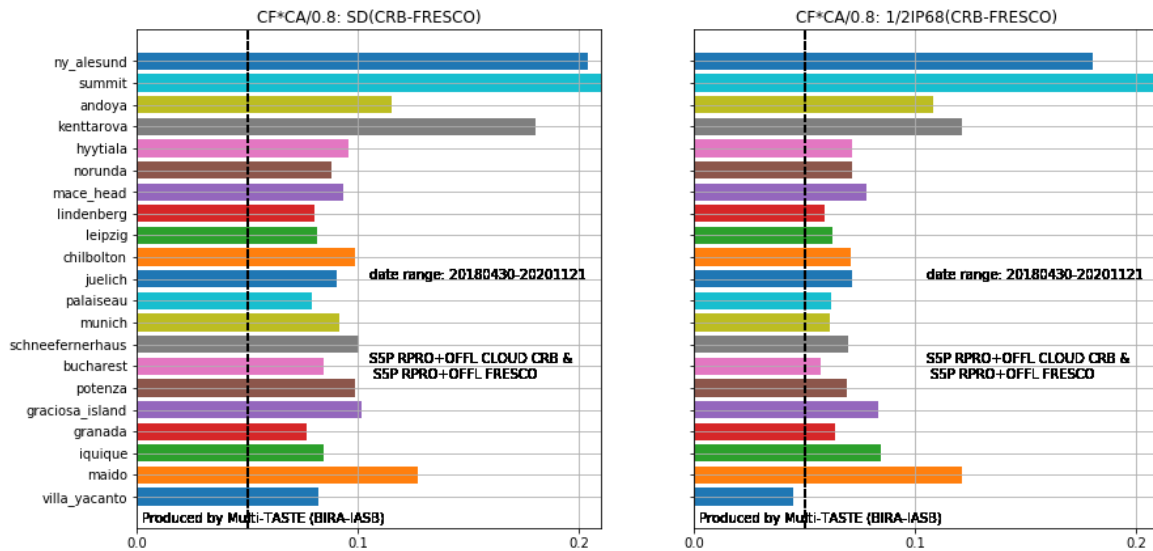


Figure 86. Left. Standard deviation S5p CLOUD CRB cloud fraction minus S5p FRESCO cloud fraction, after both have been rescaled to a cloud albedo of 0.8. Right: same but for the $\frac{1}{2}$ IP68. Note that the dispersion at the site Summit is off-scale and extends to ~ 0.6 km. Date range: 2018/04/30-2020/08/24.

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 Compernelle et al. (2020), Figure S1.

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

11.3.4.1 Bias

Comparison with CLOUDNET cloud top height and cloud height

L2_CLOUD CAL cloud top height is generally below the CLOUDNET cloud top height. A typical case is provided for the CLOUDNET site at Jülich (Figure 87, Figure 88). The monthly mean S5P CAL CTH generally follows the trend of the CLOUDNET cloud top height (Figure 87). This is corroborated by the high Pearson correlation coefficient of 0.78 (Figure 88). In absolute scale terms, the underestimation is higher for high clouds (Figure 88). We note that the TROPOMI pixel size reduction (along track reduction from 7 to 5.5 km at 2019/08/06) does not seem to impact the comparison (Figure 87). The UPAS version upgrade to 2.1.3 seems to lead to an increased bias, but more data is needed to conclude. Note that a more visible impact is seen at the site Ny-Alesund (see section 11.4).

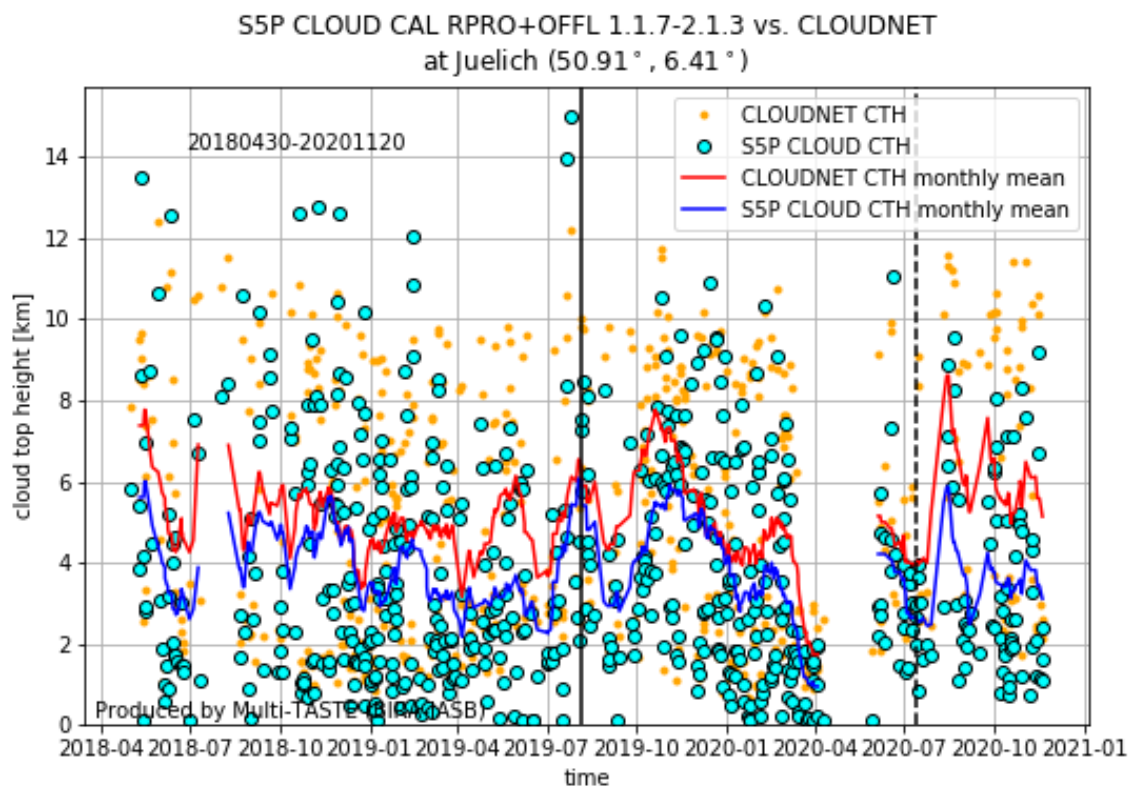


Figure 87: Time series of S5P CLOUD CAL (RPRO and OFFL, processor version 01.01.07-02.01.03) CTH vs. CLOUDNET CTH at Jülich. The monthly mean of both is also provided. The moment of the switch in TROPOMI pixel size (along track reduction from 7 to 5.5 km) is identified by the bold vertical line, the UPAS processor upgrade to v02.01.03 by the dashed line. Sensing time range is indicated on the figure.

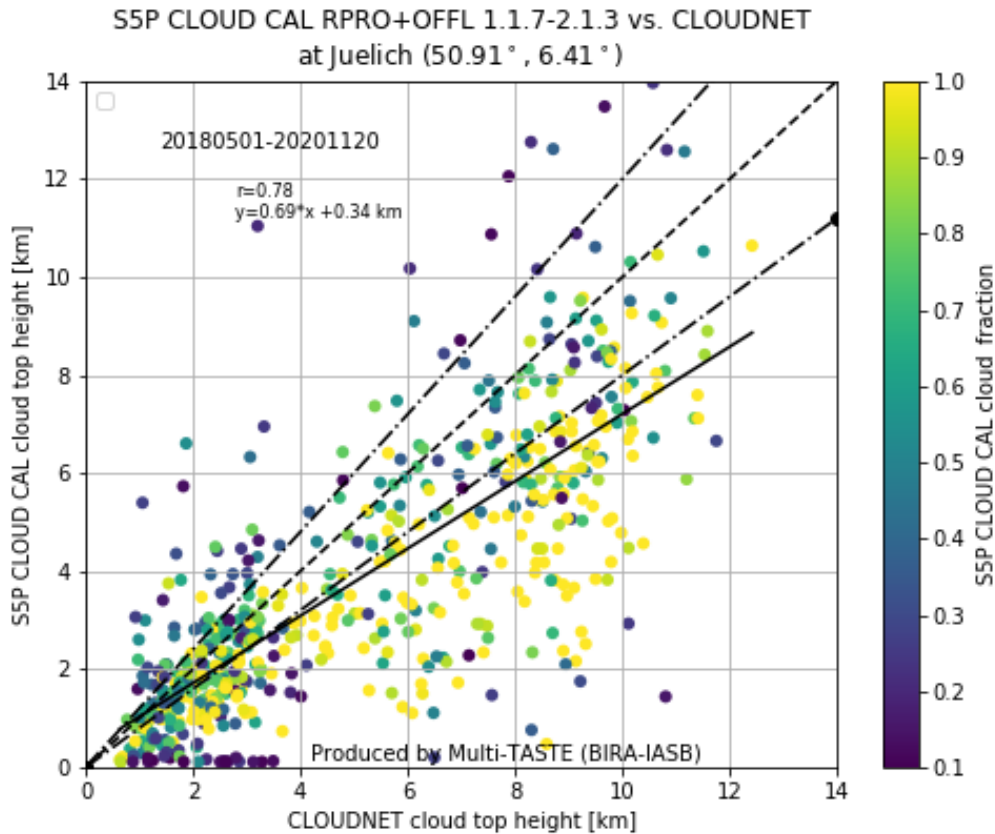


Figure 88: Correlation plot of S5P L2_CLOUD CAL (RPRO+OFFL, processor version 01.01.07-02.01.03) CTH vs. CLOUDNET CTH at Jülich. The colour indicates the S5P L2_CLOUD cloud fraction. Dashed line is the 1:1 line and dash-dotted line the 20% bias requirement.

At most of the sites, the S5P L2_CLOUD CAL top height is lower than the CLOUDNET top height, as depicted in **Figure 89**. 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. Values of -1 km and -25% (after rounding to 5%) is obtained. The bias depends on the cloud height. In distribution plots of CLOUDNET cloud top heights two modes are typically visible (see e.g., Compennolle, 2020, Fig. 10), where the higher mode contains more ice cloud and multilayer clouds. If one considers 4 km as a rough threshold between low (<4 km) and high (>4 km) clouds, the bias (relative bias) is -2 km (-30%) for high clouds, and -0.3 km (-15%) for low clouds.

We note that in S5P L2_CLOUD ROCINN_CAL a cloud is modelled as consisting of a homogeneous single-layer liquid droplets. It is therefore reasonable that the agreement with high CLOUDNET CTH clouds, which consist mostly of ice clouds and/or are of multilayer nature is less than with the low CLOUDNET CTH clouds. We further note that since version 2.1.3, a cloud_phase flag was added to S5P L2_CLOUD and an ice cloud parameterization in the retrieval is under development and intended to be implemented in a future release.

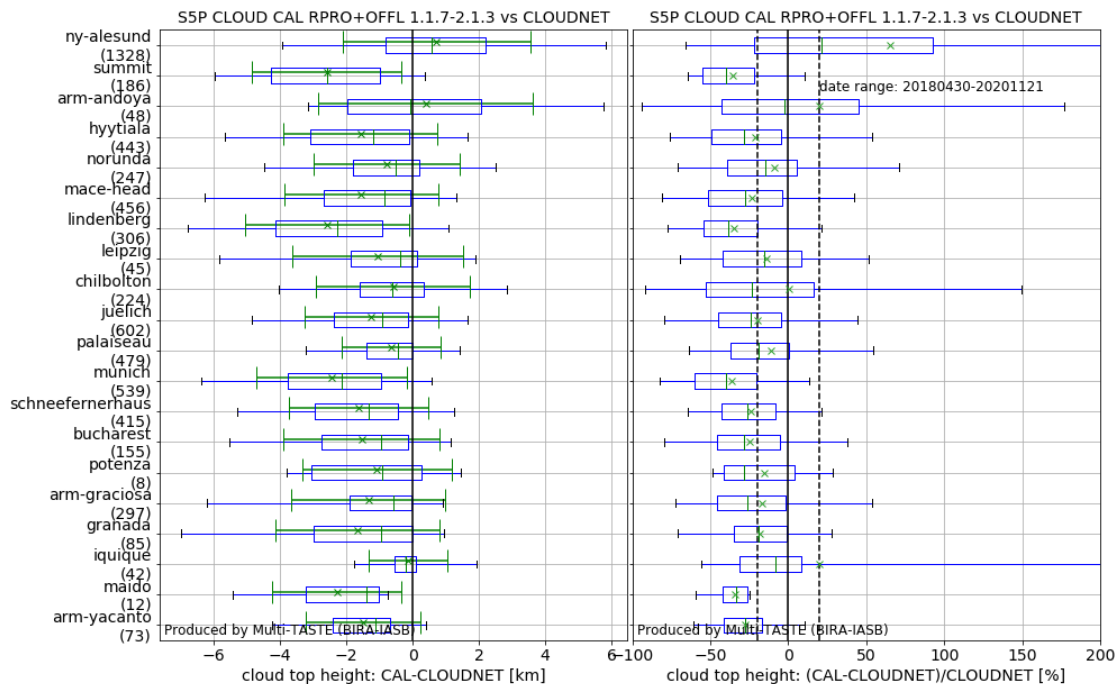


Figure 89: Left panel: Boxplots of S5P L2_CLOUD CAL CTH minus CLOUDNET CTH, per site. Boxplot conventions: box bounds are at first and third quartile. Green line is median. Whiskers are at 5 and 95 percentiles. Slightly offset green cross is mean, with green error bar $\pm$ the standard deviation. Right panel: Similar as upper panel but now for the relative difference, and without indication of the standard deviation. Sensing date range is indicated on the plot.

The mean relative difference is in about half of the cases lower than the bias upper limit requirement of 20%. Inspection of the individual cases (not shown here) lead to the following conclusions:

- The bad agreement at Summit can be ascribed to the occurrence of snow and ice (a known problem for satellite retrieval), resulting in retrieving the surface height as the cloud top height.
- The bad agreement at Schneefernerhaus is due to the specific orography; the CLOUDNET station is at a mountain, and the S5P surface altitude for co-located pixels is approximately 1 km below the CLOUDNET station. S5P CLOUD (as well as the other UPAS and also the KNMI products) takes its surface altitude from the DEM GMTED2010, but averaged within a radius of 5 km.
- The large positive bias at Ny-Ålesund is caused mainly by a cluster of low CLOUDNET CTH values, with a high retrieved CLOUD CTH and a low retrieved CLOUD CF. This site is challenging as land, sea, coast, and (depending on the season), snow/ice and sea-ice can all occur within one pixel. Note that the switch to version 2.1.3 has a big impact (see Section 11.4). Also, the high-latitude site Andoya has a bad agreement.
- The cause of the exceptionally large negative bias for Munich and Lindenberg is probably due to the relatively high proportion of higher CLOUDNET cloud heights, and therefore a high proportion of ice clouds, in the collocated data (Compennolle et al., 2020). The ROCINN algorithm works better for liquid clouds.

CLOUD CRB cloud height is generally below the CLOUDNET cloud height. A typical case is provided for the CLOUDNET site at Jülich (Figure 90, Figure 91). The monthly mean S5P CRB CH generally follows the trend of the CLOUDNET cloud mean height (Figure 90). This is corroborated by the high Pearson correlation coefficient of 0.79 (Figure 91). We note that the TROPOMI pixel size reduction (along track reduction from 7 to 5.5 km at 2019/08/06) does not seem to impact the comparison (Figure 87). The UPAS

version upgrade to 2.1.3 seems to lead to an increased bias, but more data is needed to conclude. Note that a more visible impact is seen at the site Ny-Alesund (see section 11.4).

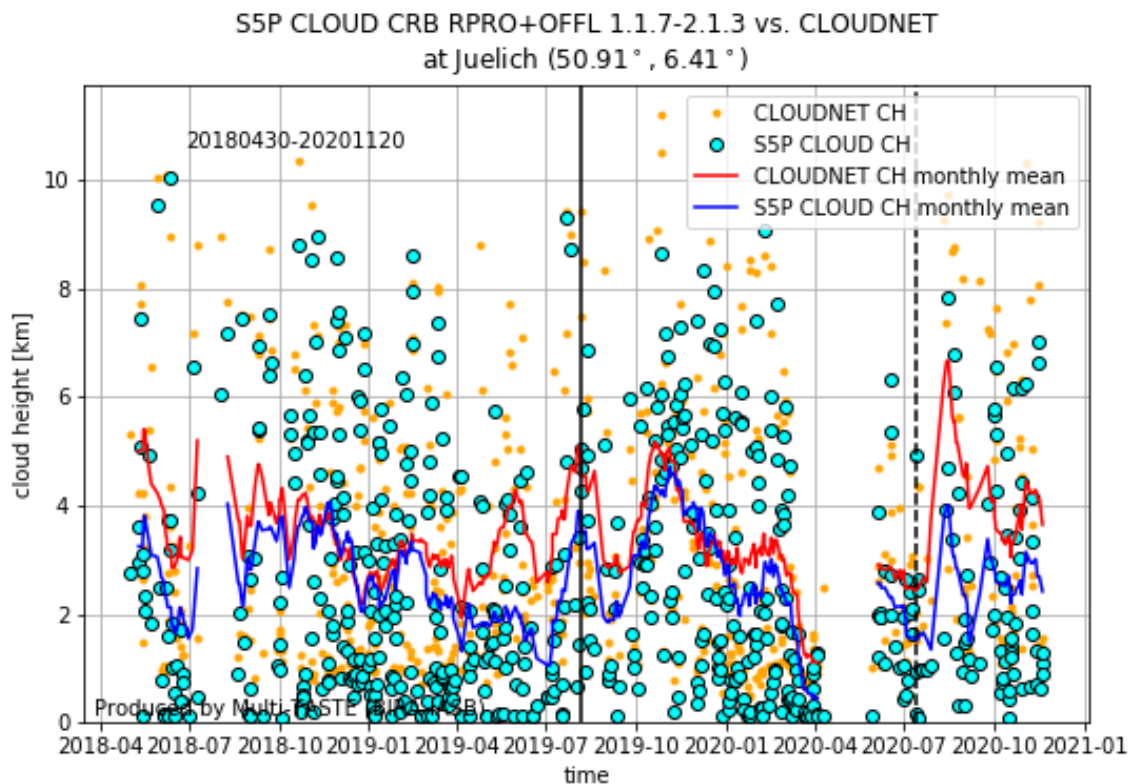


Figure 90: Time series of S5P L2_CLOUD CRB (RPRO and OFFL, processor version 01.01.07-02.01.03) CH vs. CLOUDNET CH at Jülich. The monthly mean of both is also provided. The moment of the switch in TROPOMI pixel size (along track reduction from 7 to 5.5 km) is identified by the bold vertical line, the UPAS processor upgrade to 02.01.03 by the dashed line. Sensing time range is indicated on the figure.

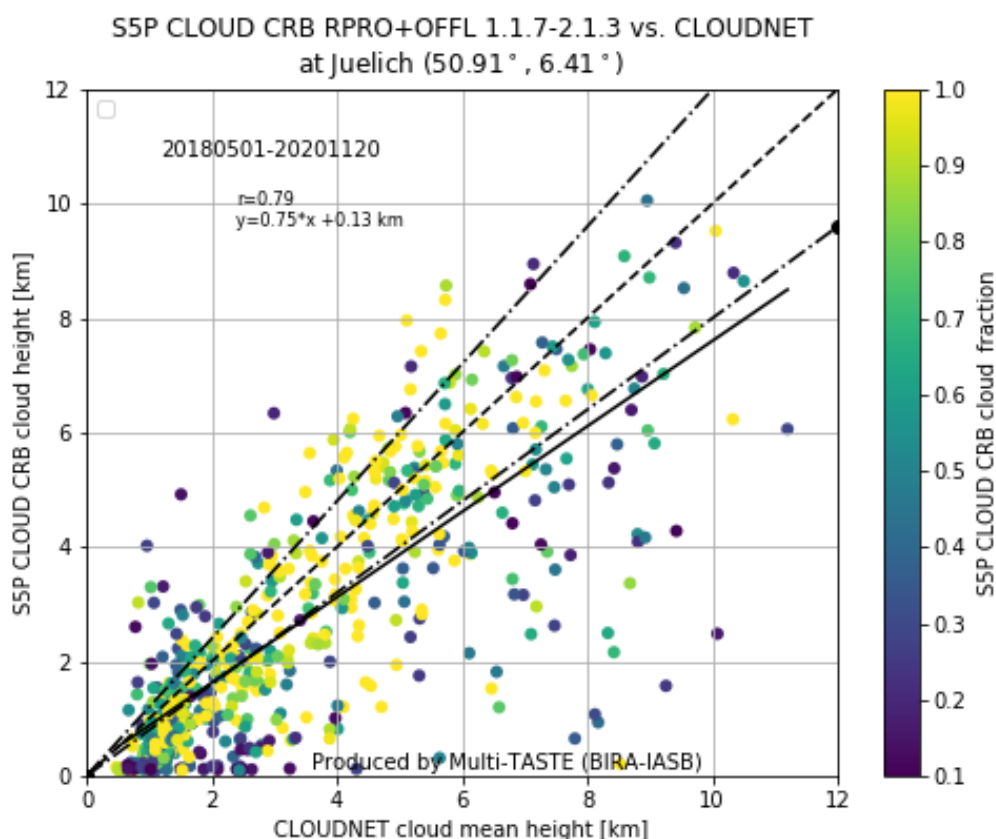


Figure 91: Correlation plot of S5P L2_CLOUD CRB (RPRO and OFFL, processor version 01.01.07-02.01.03) CH vs. CLOUDNET cloud mean height at Jülich. The colour indicates the S5P CLOUD cloud fraction. Dashed line is the 1:1 line and dash-dotted line the 20% bias requirement. Sensing time range is indicated on the figure.

To obtain an overall value for the bias and the relative bias, we calculate the median over all per-site medians of the difference and of the relative difference. Values of -0.6 km and 30% (rounded to 5%) is obtained.

At most of the sites the S5P L2_CLOUD CRB CH is lower than the CLOUDNET mean height. At most sites the bias is below 750 m. The 20% limit is exceeded in more than half of the cases. The outliers at Schneefernerhaus and Summit are due to orography and snow/ice cover (see the discussion for CAL CTH bias). Exceptions Munich, Lindenberg are probably due to the higher proportion of high (ice) clouds in the collocated data. Roughly similar biases are seen for the S5P support product FRESCO vs. CLOUDNET, indicating that most discrepancies are not specific to a particular cloud retrieval algorithm. However, note that the comparison of CLOUD CRB CH vs CLOUDNET CH at Iquique is impacted by outlying values (as can be seen from the long tail in the boxplot in Figure 92) while this is not the case for FRESCO vs CLOUDNET. We return to this in Section 11.3.4.3.

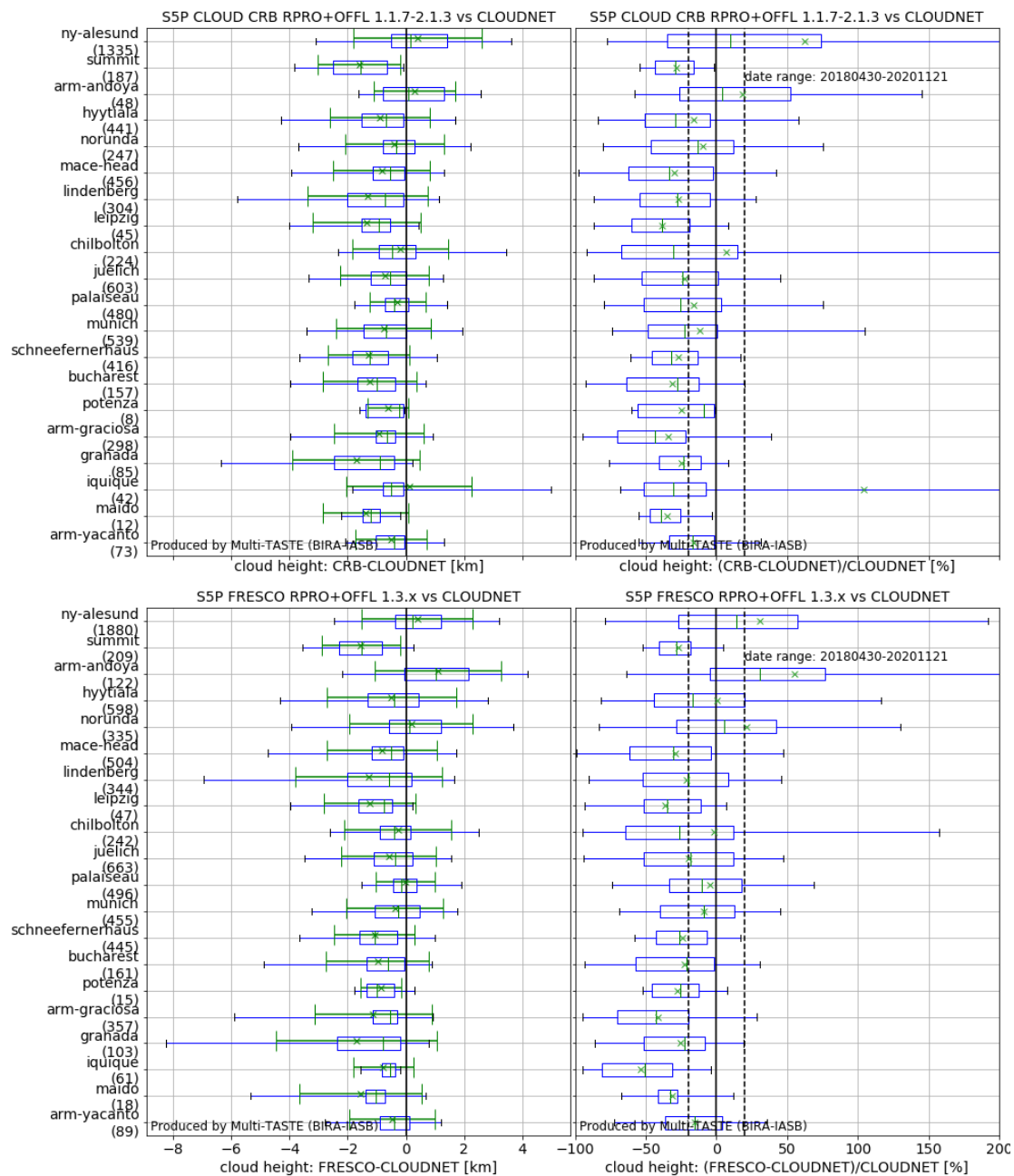


Figure 92: Upper panel: Boxplots of S5P L2_CLOUD CRB RPRO+OFFL 01.01.07-02.01.03 CH minus CLOUDNET CH (upper left) and of the relative difference (upper right), per site. Lower panel: The same but now for the S5P Support product FRESKO RPRO 01.03.02 + OFFL 01.03.00-02. The same conventions as for **Figure 89** apply. Sensing time range is indicated on the figure.

Comparison with alternative S5p cloud height retrievals

S5p CLOUD CRB CH was compared with S5p FRESKO CH, over the CLOUDNET sites. At most sites, a small bias is encountered between 0 and -0.2 km (mean and median difference) or ~-10% (median normed relative difference). This is within the 20% bias requirement. Clear exceptions (CRB higher than FRESKO) occur at the sites Iquique, Ny-Ålesund and Summit.

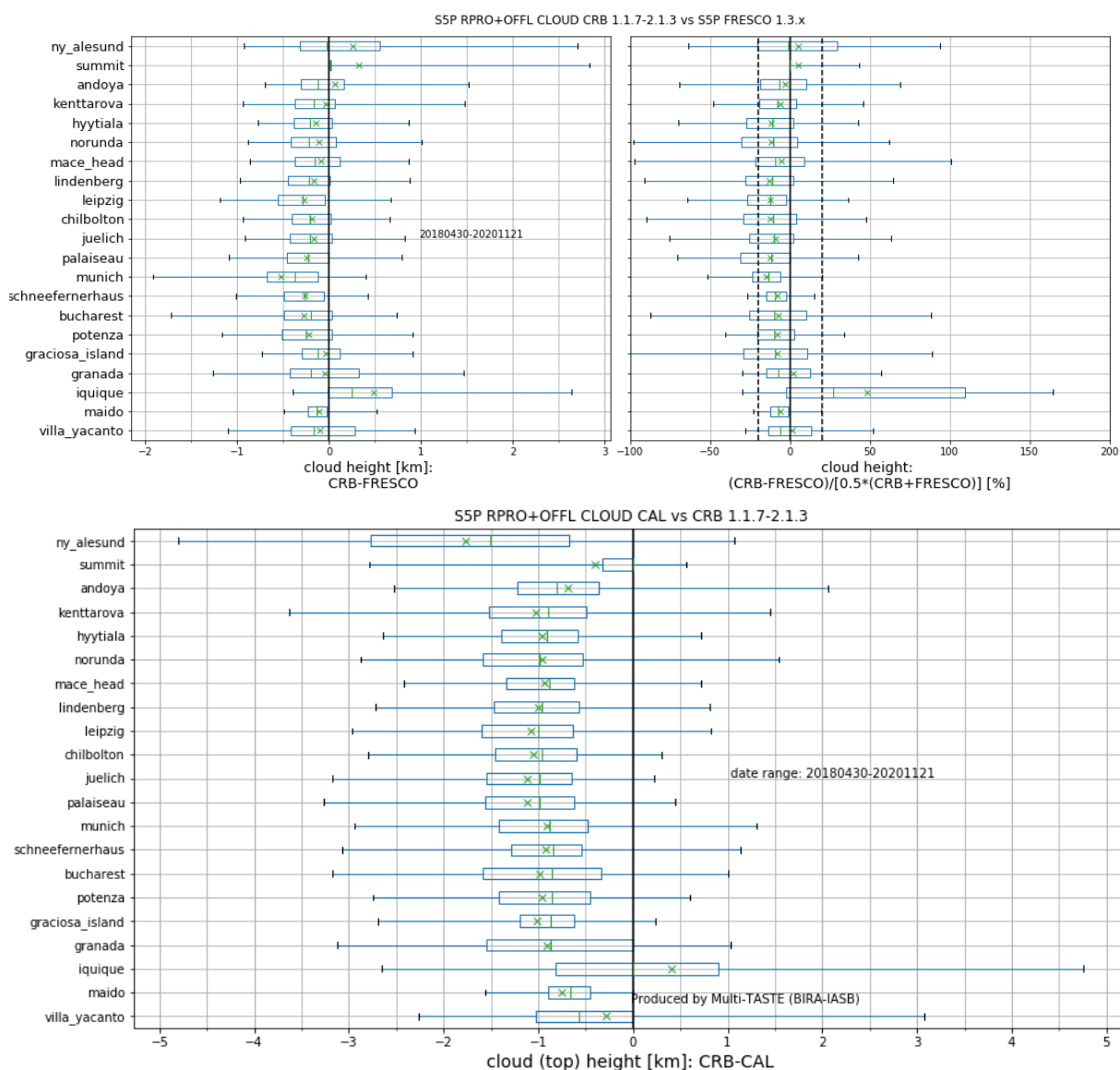


Figure 93: Top. Boxplots of S5p CLOUD CRB CH minus S5p FRESCO CH (left) and of the normed difference (right). Bottom. Boxplots of S5p CLOUD CRB CH minus S5p CLOUD CAL CTH.

S5p CLOUD CRB CH was also compared with Sp5 CLOUD CAL CTH. As expected, CLOUD CRB CH is lower than CLOUD CAL CTH; the mean and median difference is typically -1 km. Again, exceptions occur at Iquique, Ny-Ålesund and Summit. At Iquique, the mean difference is positive.

Figure 84 presents zonal means of S5P CLOUD CRB CH and S5P FRESCO CH at 2020-02-29. Both have a similar variation, but with S5P FRESCO CH slightly below S5P CLOUD CRB CH.

Comparison with MODIS, NPP VIIRS and OMI OMCLDO2

S5P CLOUD CAL CTH version 1 was compared to pre-production VIIRS data (Compennolle et al., 2020). A negative bias (mean difference) in the cloud top height (CTH) (-1.6 km) has been found. Note however that S5P L2_CLOUD and VIIRS capture the same CTH mode at 1.8 km (Fig. 8 of Compennolle et al., 2020).

A comparison of S5P L2_CLOUD with MODIS using commissioning phase data also showed a negative CTH bias (-1 km between -60°S and +60°S).

In **Figure 84**, the latitudinal variation of zonal means at 2020-02-29 for S5P CLOUD CAL CTH, S5P CLOUD CRB CH, OMI OMCLDO2 CH and MODIS CTH are compared. Similar variations are found, with S5P CLOUD CRB CH < S5P CLOUD CAL CTH (as expected), OMI OMCLDO2 CH comparable to S5P CLOUD CAL CTH, and MODIS CTH slightly higher.

11.3.4.2 Dispersion

Comparison with CLOUDNET cloud top height and cloud height

From **Figure 89** it can be inferred that the comparison spread (expressed as standard deviation) of S5P CLOUD CAL CTH vs. CLOUDNET CTH, exceeds the upper limit for random error dispersion (500 m). However, also CLOUDNET CTH random error, and comparison error, contribute to the comparison spread, and these contributions have not been quantified yet. The same holds for S5P CLOUD CRB CH vs. CLOUDNET CH (**Figure 92**). Roughly similar dispersions are seen for the S5P support product FRESCO vs. CLOUDNET, indicating that the discrepancies are not specific to a particular cloud retrieval algorithm. Overall values of dispersion are obtained by taking the median over all per-site $\frac{1}{2}$ IP68 values. This gives an overall dispersion of 2 km for CLOUD CAL CTH vs CLOUDNET CTH, and 1 km for CLOUD CRB CH vs CLOUDNET CH and FRESCO CH vs CLOUDNET CH.

Distinguishing between high (CLOUDNET CTH >4km) and low (CLOUDNET CTH <4 km) clouds, the dispersion of CLOUD CAL CTH vs CLOUDNET CTH is 2.1 km for high clouds and 0.8 km for low clouds.

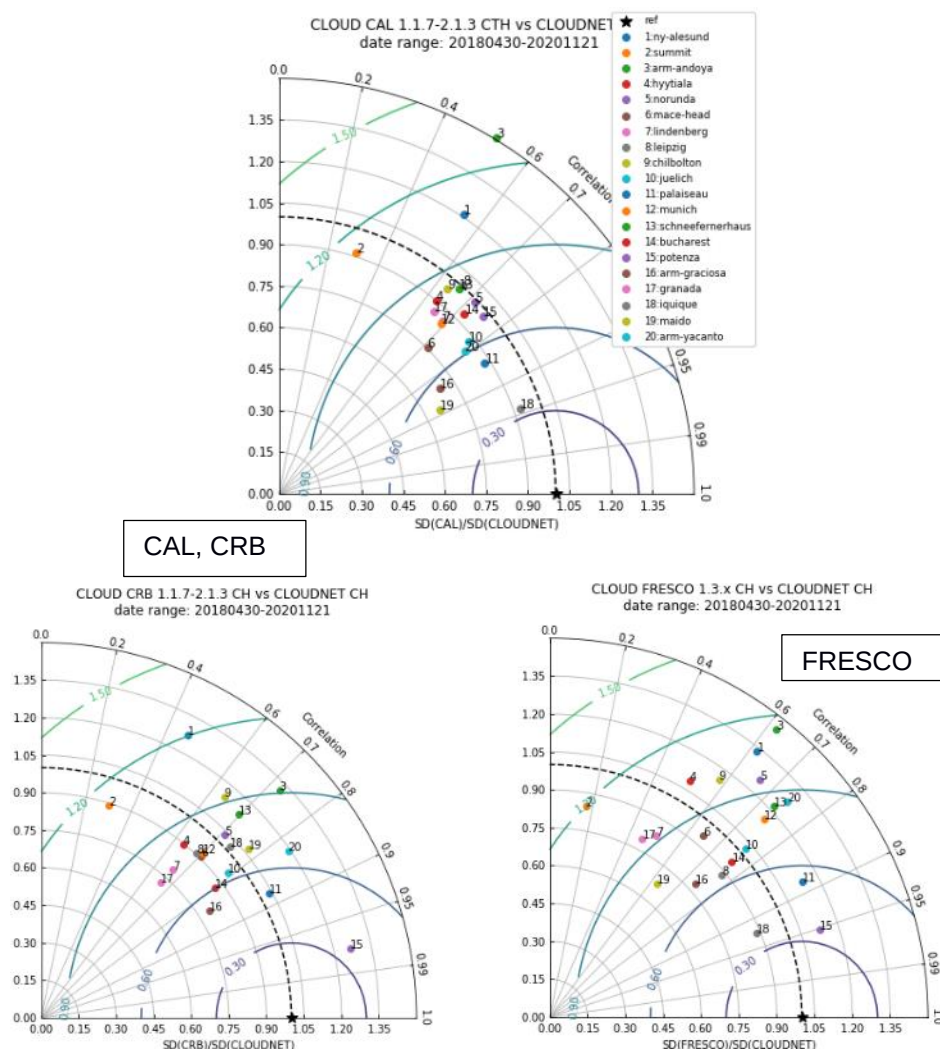


Figure 94: Taylor diagrams of CLOUD CAL vs CLOUDNET CTH, CLOUD CRB vs CLOUDNET CH and FRESCO CH vs CLOUDNET CH. The radius represents $SD(S5p)/SD(CLOUDNET)$, the angle represents correlation and the circular contour lines represents the relative dispersion $SD(S5p-CLOUDNET)/SD(CLOUDNET)$.

Figure 94 presents Taylor diagrams of CLOUD CAL CTH vs CLOUDNET CTH, CLOUD CRB vs CLOUDNET CH and FRESCO CH vs CLOUDNET CH. High relative dispersion and low correlation is found for the sites Summit, Ny-Ålesund, Andoya, Chilbolton, and specifically for FRESCO, Hyttiala; the latter could be due to a different treatment of snow-ice cases. For 16 out of 20 cases, the relative dispersion $SD(S5p-CLOUDNET)/SD(CLOUDNET)$ is larger for S5P FRESCO than for S5P CLOUD CRB. For the site Iquique, a low relative dispersion and high correlation is found with CLOUD CAL and FRESCO, but much lower with CLOUD CRB.

Comparison with alternative S5p cloud height retrievals

The comparison of S5p CLOUD CRB CH vs S5p FRESCO CH reveals a low standard deviation of the differences at most sites of 0.6-0.7 km (**Figure 95**, left). Exceptions are Iquique, Granada, Summit, Munich, Kenttarova and Ny-Alesund where the dispersion is considerably higher. The more robust dispersion measure $\frac{1}{2}$ IP68 is 0.3-0.4 km at most sites, which is below the dispersion requirement of 0.5 km.

The comparison of S5p CLOUD CRB CH vs S5p CLOUD CAL CTH (**Figure 93**, bottom) reveals a dispersion of the differences at most sites of ~ 1.2 km. Exceptions are Iquique and Ny-Alesund where the dispersion is considerably higher.

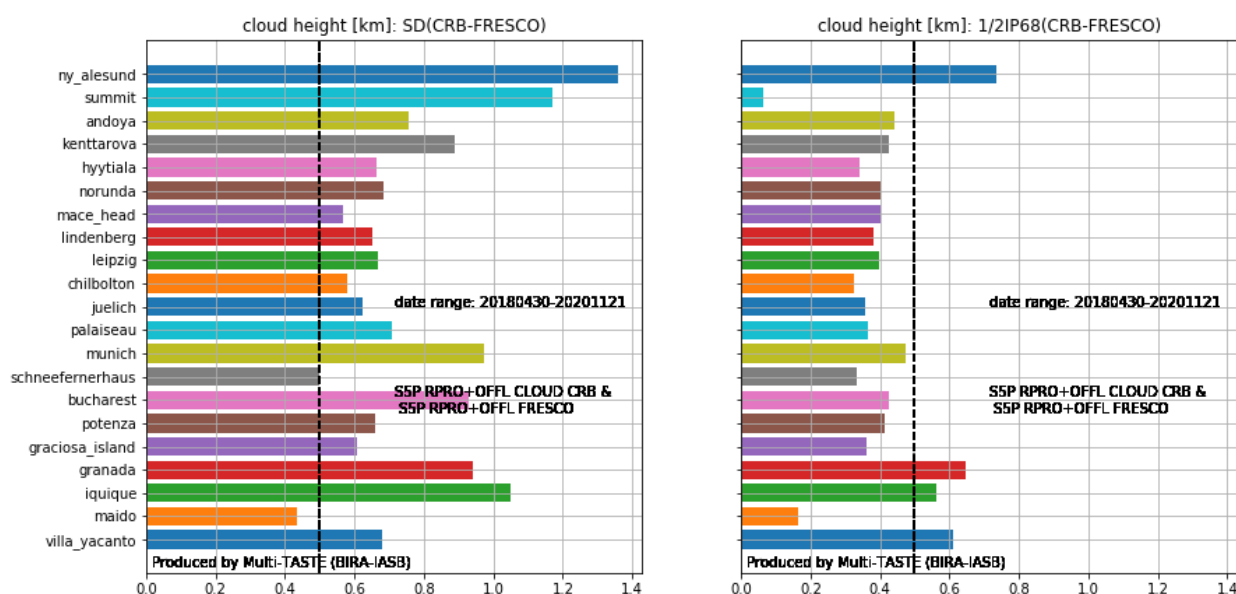


Figure 95: Standard deviation (left) and $\frac{1}{2}$ IP68 (right) of S5p CLOUD CRB CH minus S5p FRESCO CH. Sensing date range is indicated on the figure.

Comparison with NPP VIIRS

S5P L2_CLOUD CAL CTH version 1 shows good correlation with VIIRS CTH: Pearson coefficient $R = 0.74$ for continental clouds and 0.86 for marine clouds (see Fig. S9 of Compernelle et al., 2020).

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, higher discrepancies are encountered for CLOUD CRB CH – FRESCO CH (**Figure 93**, top), CLOUD CRB CH – CLOUD CAL CTH (**Figure 93**, bottom), and CLOUD CRB CH – CLOUDNET (**Figure 92**). 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. **Figure 96** demonstrates this for CLOUD CRB vs CLOUDNET. On the other hand, it should be remarked that the remaining CRB data points in the 0-2km x 0-2km box of **Figure 96** (top) are closer to the 1:1 line than the FRESCO data points in the 0-2km x 0-2km box of **Figure 96** (bottom). Note that for this site there is only collocated data up to 2019/01/09, therefore only CLOUD CRB version 1 is involved here.

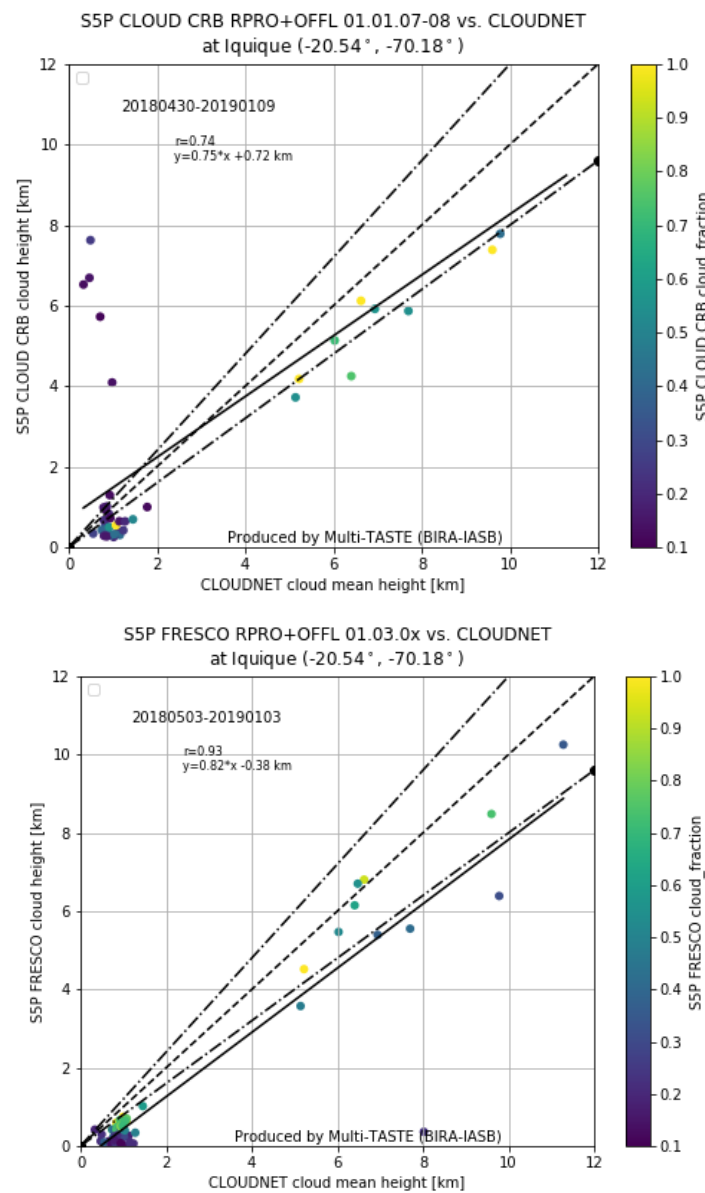


Figure 96. Top. S5p CLOUD CRB CH vs CLOUDNET CH correlation plot. Bottom. S5p FRESCO CH vs CLOUDNET CH correlation plot.

Comparison with alternative S5p cloud height retrievals

Figure 97 presents the CLOUD CRB-FRESCO cloud height difference in function of CRB cloud fraction, at Juelich. The dispersion between S5p CLOUD CRB and S5p FRESCO increases with decreasing cloud fraction. At low CF, CLOUD CRB CH tends to be higher than FRESCO CH, while the opposite is true at high CF. This is a recurrent feature at several sites.

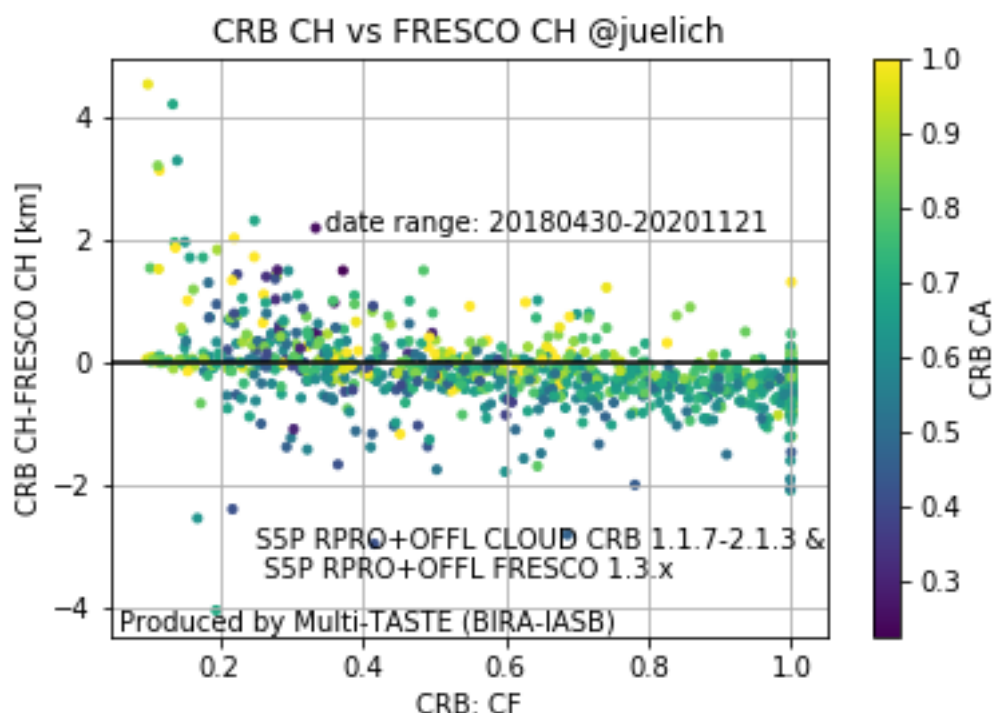


Figure 97: Cloud height difference between S5p CLOUD CRB and S5p FRESCO, in function of CLOUD CRB cloud fraction. Color scale is CLOUD CRB cloud albedo.

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 (Compernelle et al., 2020). 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. 3 of Compernelle et al. (2020).

11.3.5 Cloud optical thickness (L2_CLOUD CAL)

11.3.5.1 Bias

Comparison with NPP VIIRS

L2_CLOUD CAL COT version 1 was compared with pre-operational VIIRS data. A positive bias (+7.9) in the cloud optical thickness (COT) with respect to VIIRS has been found (Compernelle et al., 2020, Fig. 8).

A comparison of S5P L2_CLOUD with MODIS using commissioning phase data also showed a positive COT bias at tropical and middle latitudes (+3.8 between -60°S and +60°S).

11.3.5.2 Dispersion

Comparison with NPP VIIRS

For L2_CLOUD CAL COT version 1, a weak correlation with NPP VIIRS is seen. Pearson R = 0.48 for continental clouds and 0.66 for marine clouds (see Fig. S9 of Compernelle et al., 2020).

11.3.6 Cloud albedo (L2_CLOUD CRB)

11.3.6.1 Geographical patterns

The cloud albedo in UPAS 1.1.x may show a north-south gradient at all longitudes, with the northern hemisphere showing systematically larger values than the southern hemisphere (Compernelle et al., 2020). This issue is reduced in UPAS 2 (see Fig. S2 of Compernelle et al., 2020).

11.4 Preliminary validation of UPAS processor upgrade to version 2.1.3

The UPAS processor, of which L2_CLOUD is part of, did undergo a major upgrade in August 2020, from 1.1.8 to 2.1.3. This occurred at 2020/07/13, orbit 14239 for OFFL and at 2020/07/16, orbit 14285 for NRTI. Here we report a preliminary investigation of the impact of this upgrade.

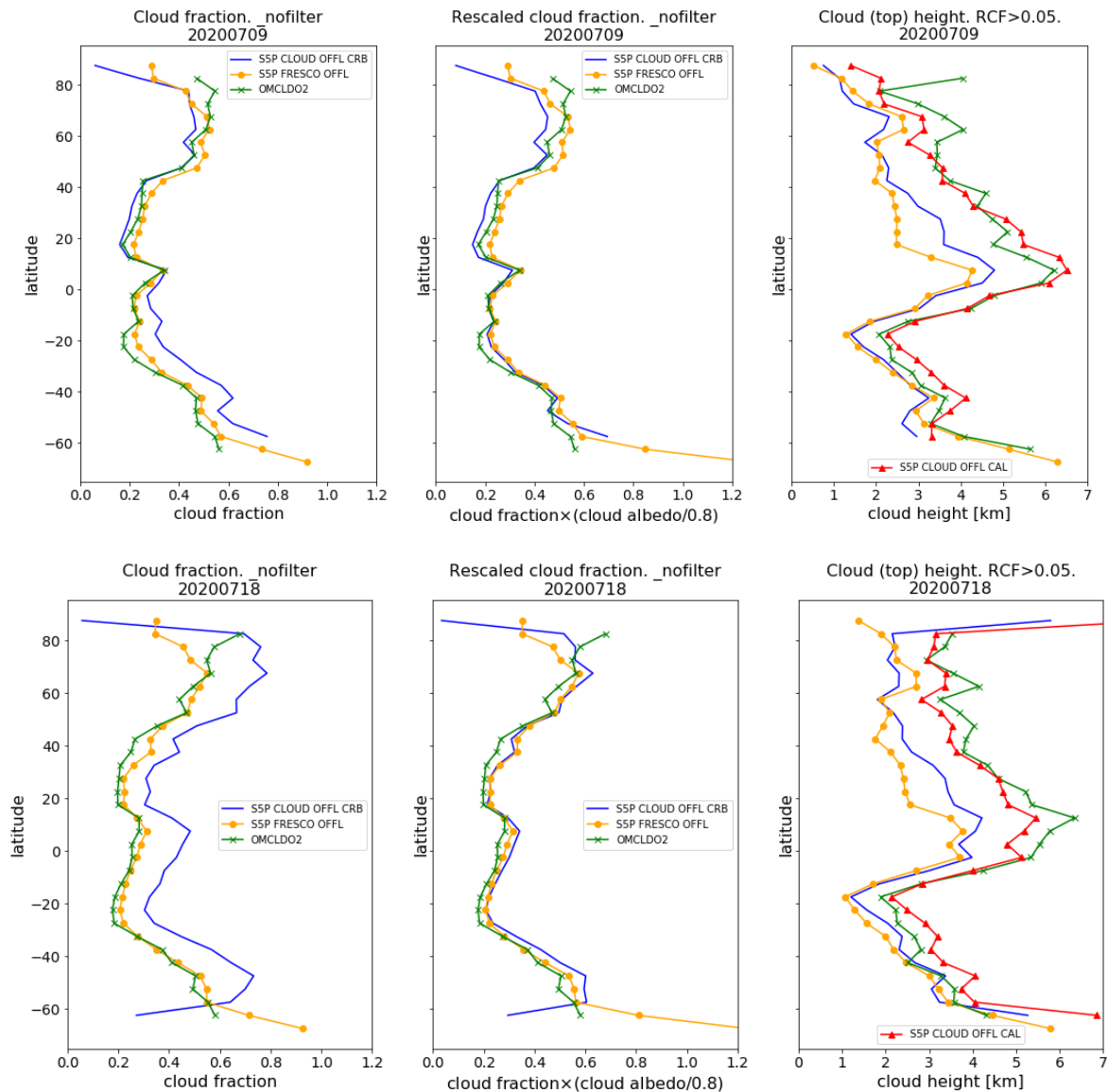


Figure 98: Latitudinal variation of zonal means of cloud fraction (left), scaled cloud fraction (middle) and cloud (top) height (right) of the cloud products S5p OFFL CLOUD (CAL and CRB), S5p OFFL FRESCO and Aura/OMI OMCLDO2. Top: at 2020/07/09 (UPAS version = 1.1.8). Bottom: at 2020/07/18 (UPAS version = 2.1.3).

Figure 98 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, which is before and after the UPAS version respectively.

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

- The different cloud products exhibit similar latitudinal variation also after the UPAS upgrade.
- 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) are better comparable quantities. The scaled radiometric cloud fraction of CLOUD CRB is close to that of S5p FRESCO and OMCLDO2, although it seems to be slightly above OMCLDO2 after the UPAS upgrade. Note that the fact that scaled radiometric cloud fractions still match, implies that CRB cloud albedo is decreased with the UPAS upgrade.
- The ordering in cloud (top) height is the same before and after the UPAS upgrade: FRESCO CH < CLOUD CRB CH < CLOUD CAL CTH ≈ OMCLDO2 CH.

For most sites, time series of S5P CLOUD CAL CTH vs CLOUDNET, CTH and of S5P CLOUD CRB CH vs CLOUDNET CH, have not yet much data after 2020/07/13. But for the northern site Ny-Alesund there are clearly much less high-lying S5P data points after the version upgrade. While before the version switch S5P CLOUD C(T)H was typically above that of CLOUDNET at this site, it now drops below.

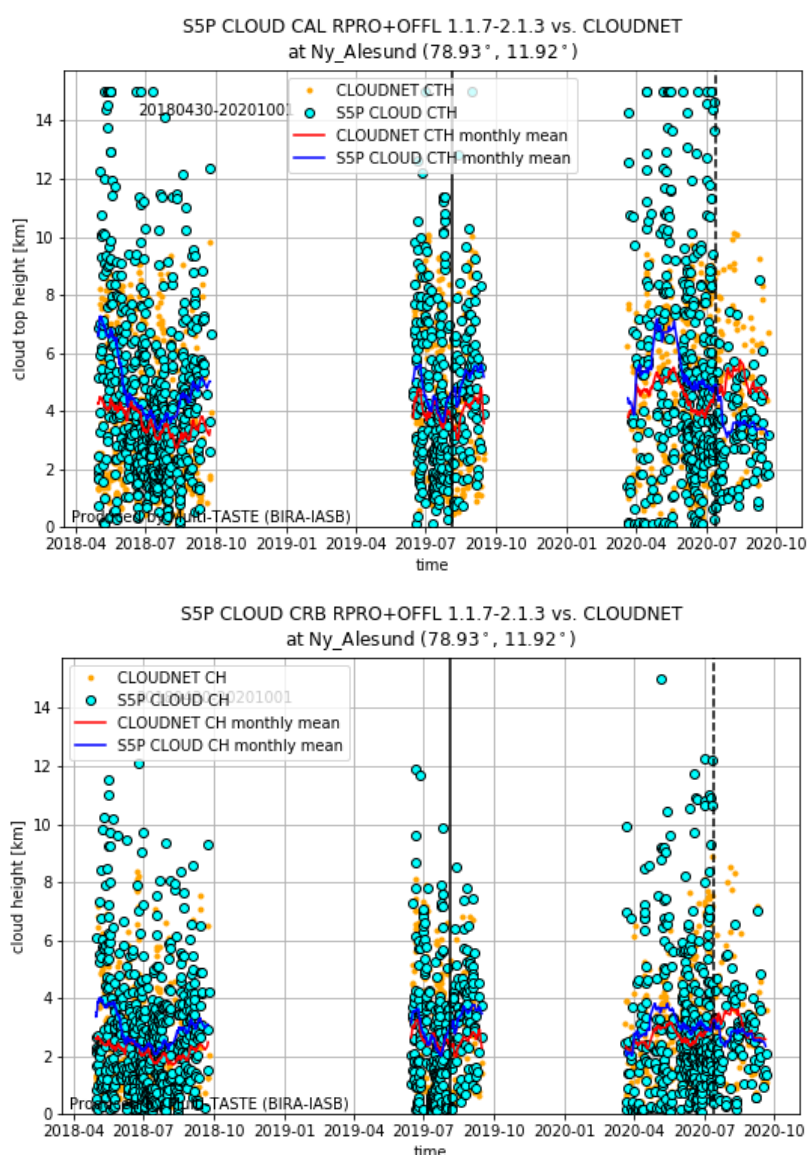


Figure 99: Top. Time series of S5P CLOUD CAL (RPRO and OFFL, processor version 01.01.07-02.01.03) CTH vs. CLOUDNET CTH at Ny-Alesund. The monthly mean of both is also provided. Sensing time range is indicated on the figure. The vertical black solid line indicates the moment of the TROPOMI pixel size switch at 2019/08/06. The vertical black dashed line indicates the version change to UPAS 02.01.03. Bottom. Same but for S5P CLOUD CRB CH vs CLOUDNET CH.

11.5 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 100 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.

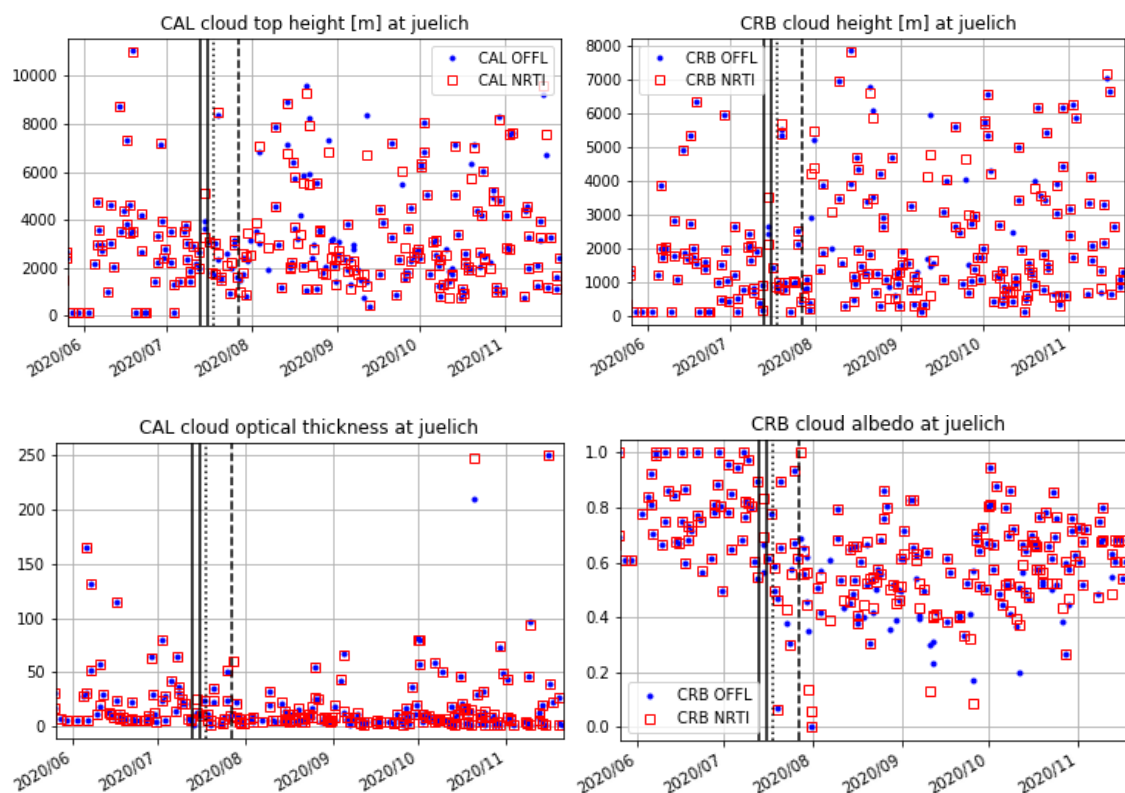


Figure 100: Time series of cloud properties of CLOUD OFFL and CLOUD NRTI at Jülich. Vertical black lines indicate the processor version change from 1.1.7 to 2.1.3; the first one for OFFL (at 2020/07/13), the second one for NRTI (at 2020/07/16).

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 0 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 10** 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 101**. 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 10 – 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

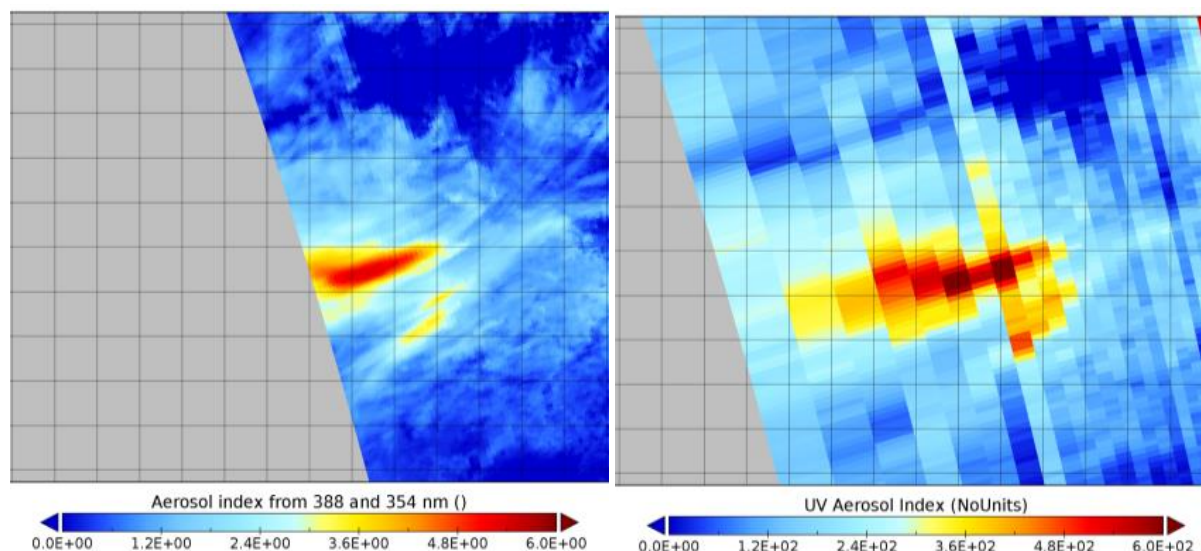
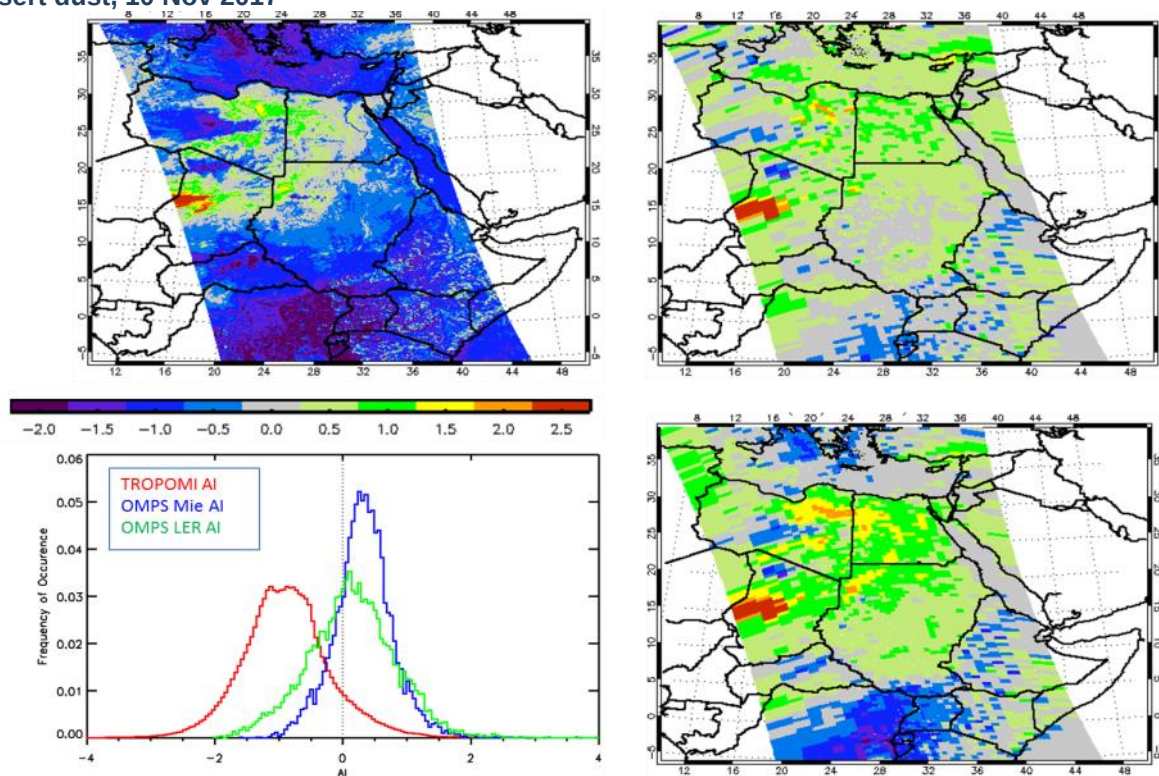


Figure 101: Comparison of S5P TROPOMI UVAI (orbit 00398, left) and OMI OMAERO UV Aerosol Index (orbit 70864, right) for Saharan dust on 10 November 2017. In general very good agreement is found (the stripes in north-south direction in the OMI data are caused by the OMI row anomaly and should be ignored).

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 102** and **Figure 103** below (courtesy of Omar Torres and Changwoo Ahn, NASA-GSFC). At the beginning of TROPOMI measurements (Nov. and Dec. 2017, **Figure 102**), 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 103**). 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 106**). 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

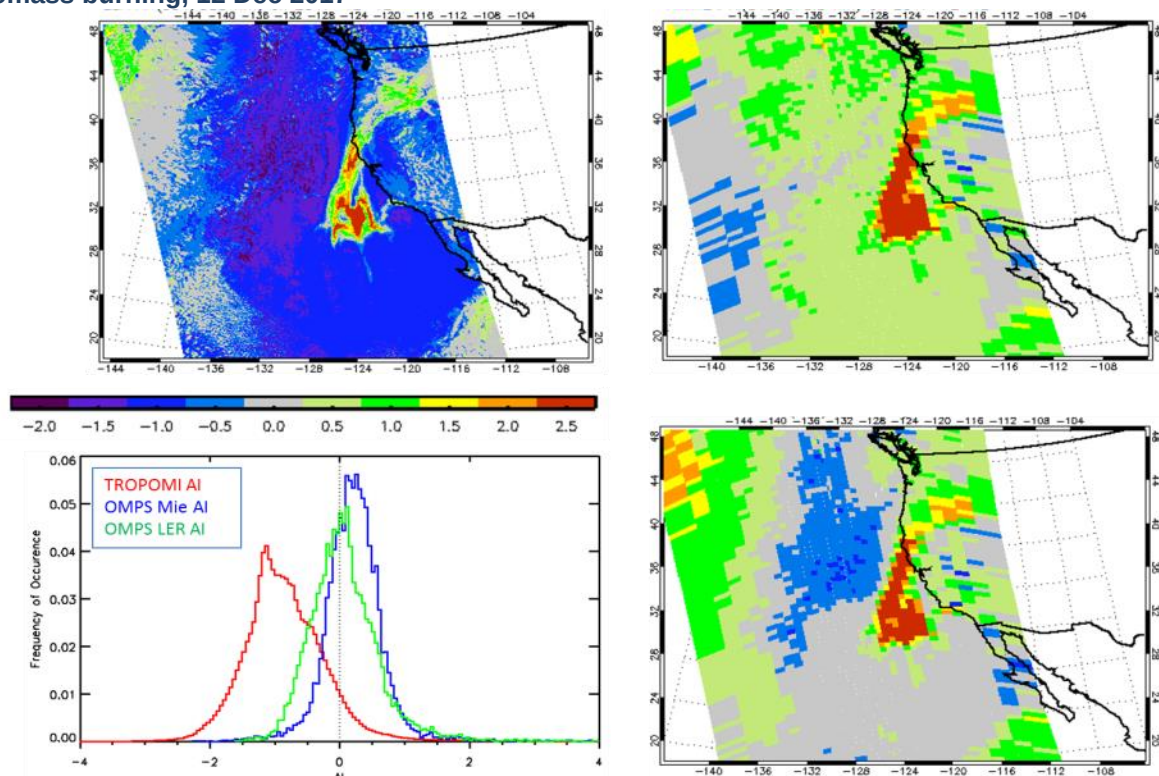


Figure 102: Comparison of UVAI from TROPOMI and OMPs for a situation with desert dust (10 Nov 2017, top) and biomass burning (12 Dec 2017, bottom). For OMPs, UVAI are calculated either assuming LER or Mie clouds. The UVAI for Mie clouds yield more consistent results. The frequency distributions indicate 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).

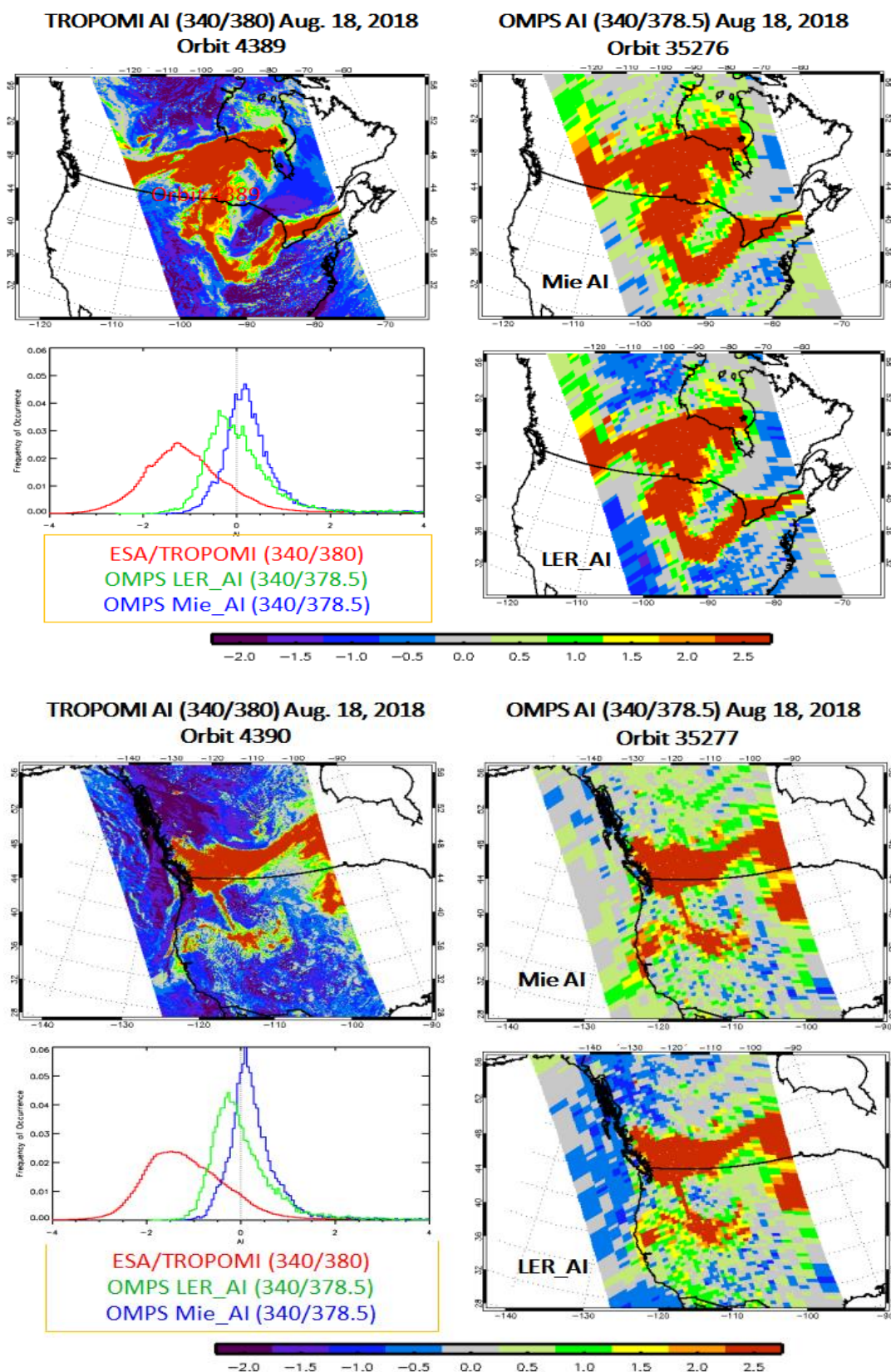


Figure 103: Comparison of UVAI from TROPOMI and OMPs for an observation of a biomass burning plume (18 Aug. 2018). For OMPS UVAI are calculated assuming either LER or Mie clouds. The UVAI for Mie clouds yields more consistent results. In contrast to early TROPOMI observations, values have systematically decreased and the spread of the UVAI values has become larger (courtesy of Omar Torres and Changwoo Ahn, NASA-GSFC).

Also comparisons with patterns from other S5P TROPOMI products are performed. **Figure 104** below shows an example of measurements of UVAI and NO₂ VCDs, for which enhanced NO₂ and UVAI are found at the same locations.

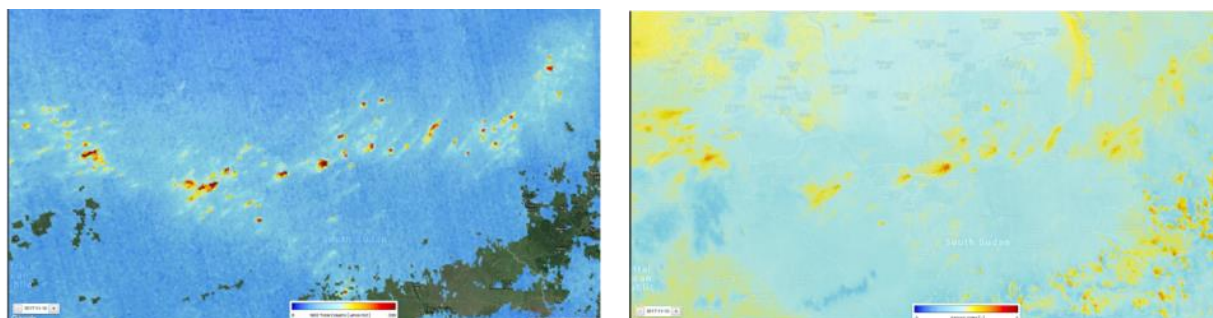


Figure 104: Comparison of NO₂ VCDs (left) and UVAI (right) obtained from S5P TROPOMI for sub-Saharan fires on 10 November 2017.

From the performed validation studies it is concluded that the L2_AER_AI UVAI from S5P TROPOMI is of very good quality and fulfilled the requirements until early 2019. The negative bias found in the S5P TROPOMI data, which continues to increase systematically is outside the bias requirements (± 1 UVAI unit) since the beginning of 2019. Here it should be noted that the bias is caused by the degradation of the level 1 irradiance data 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 10** 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 greatly 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 105** 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.

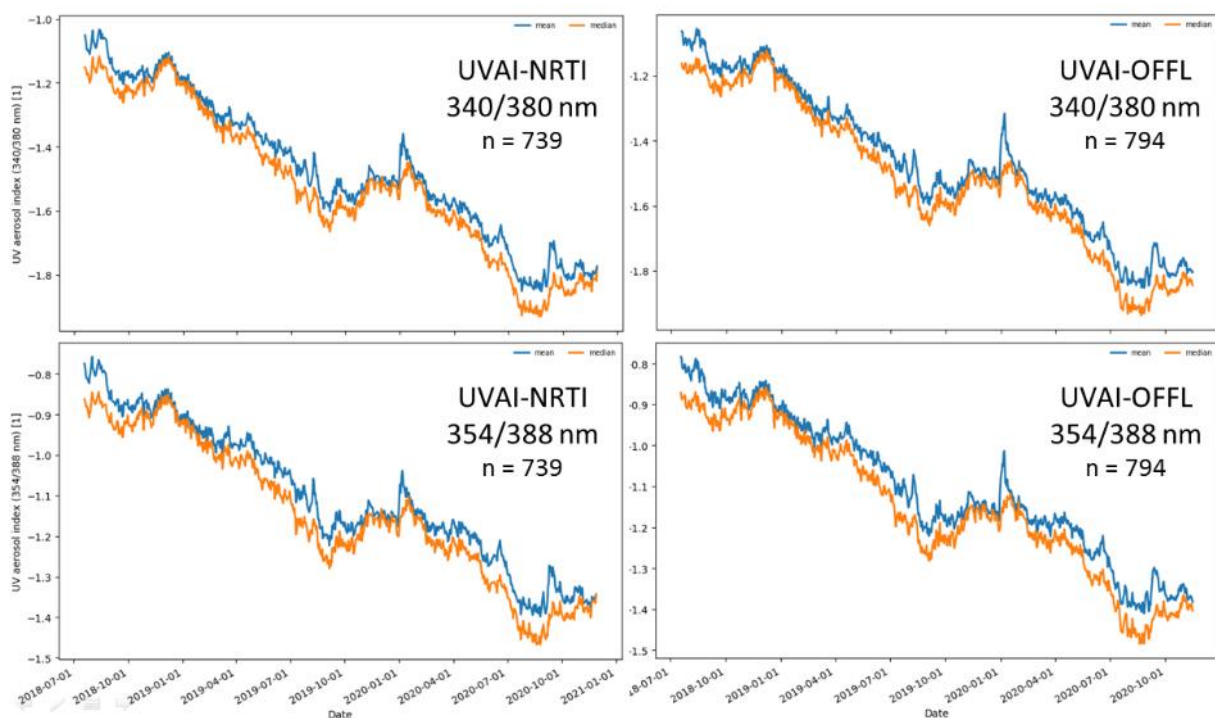


Figure 105: Comparison of the global daily mean (blue) and median (orange) for both L2_AER_AI UVAI wavelength pairs (340/380 and 354/388 nm) and for the NRTI and OFFL data streams, from 20 July 2018 through November 2020.

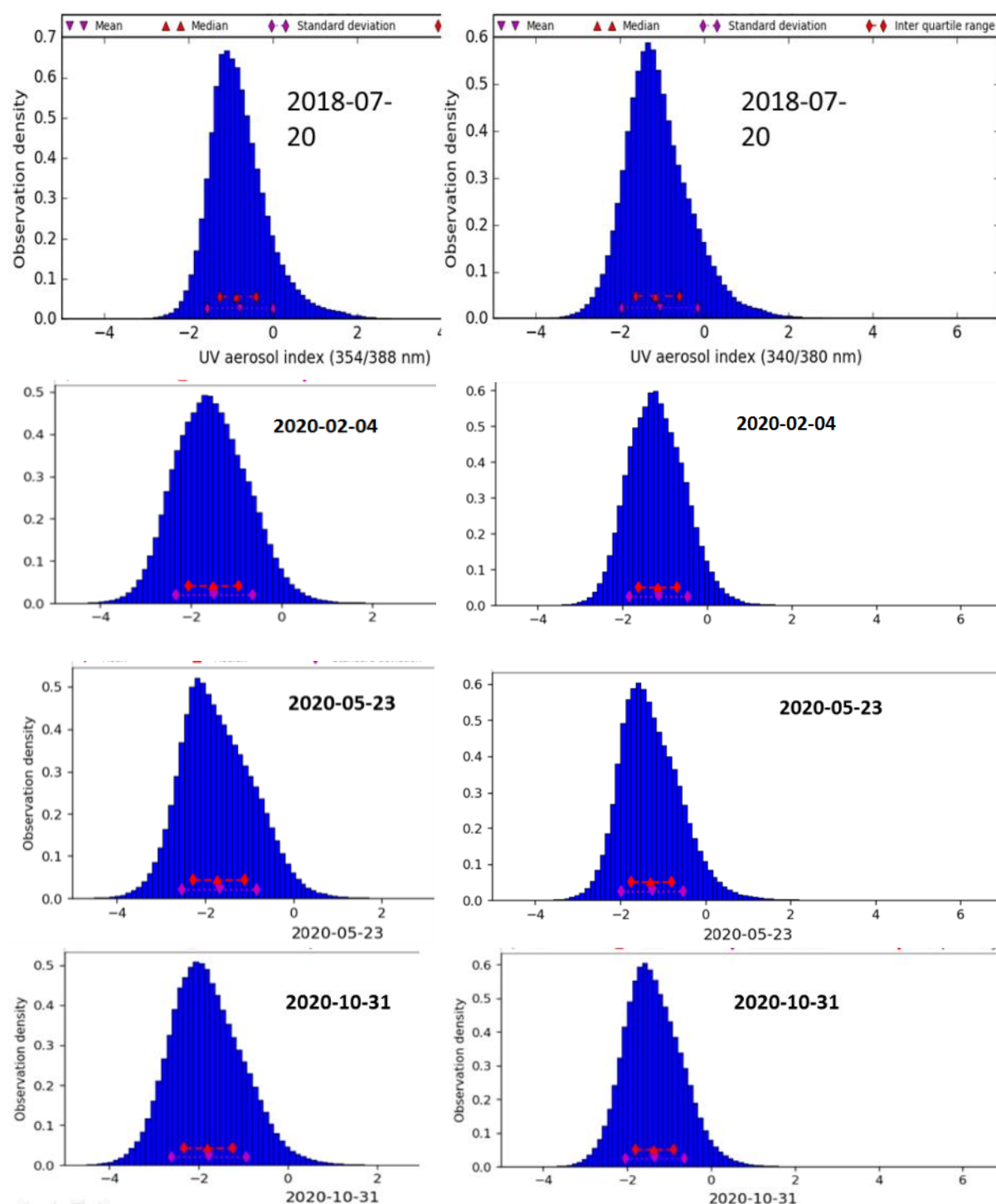


Figure 106: Comparison of the frequency distribution of the UVAI (left: 354/388nm, right: 340/380nm) for four selected days (20 July 2018, 04 February 2020, 23 May 2020, and 31 October 2020).

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 107 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.

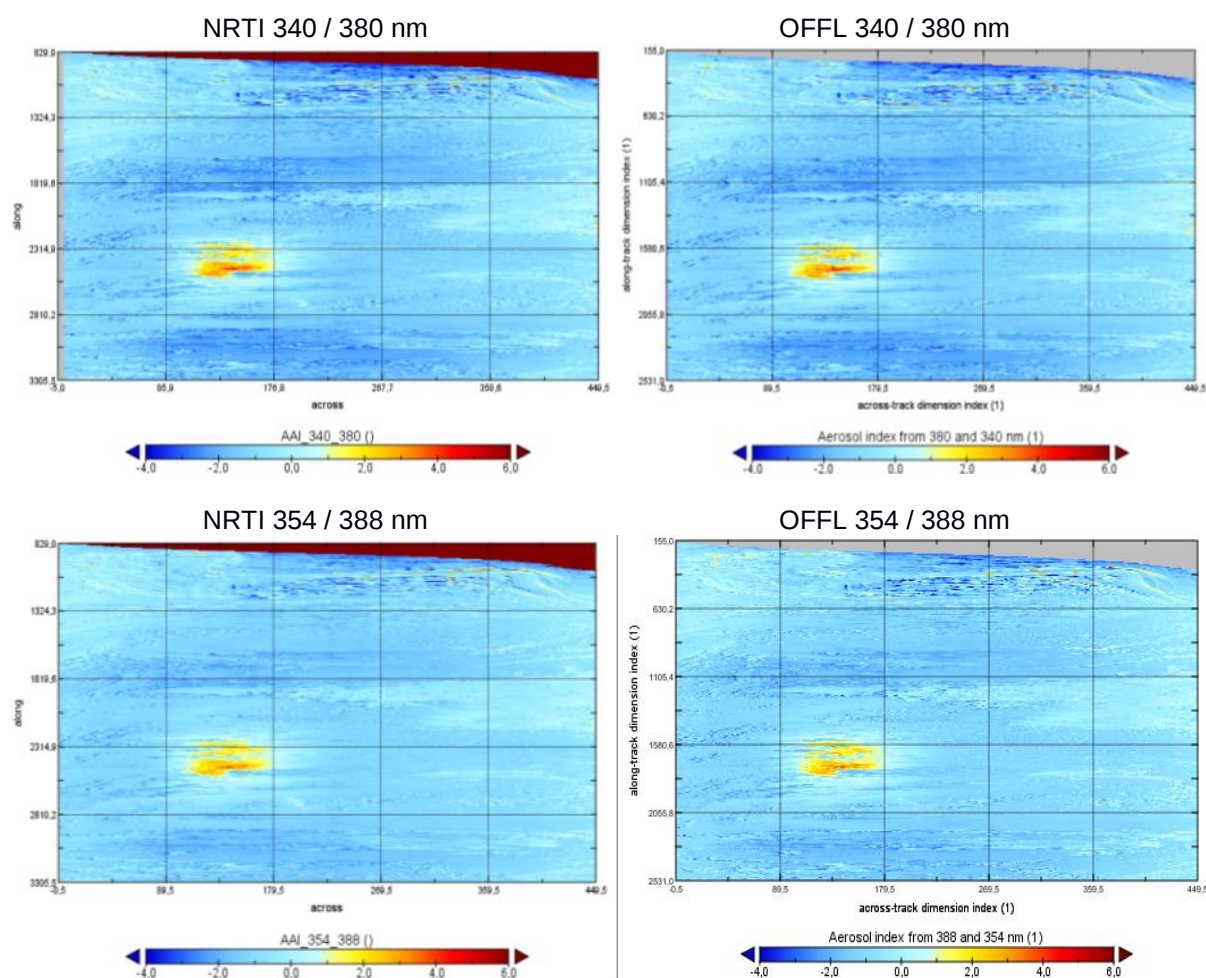


Figure 107: Comparison of the S5P TROPOMI UVAI for a selected orbit (#05033) on 3 October 2018 for the two wavelength pairs (top: 340 / 380 nm, bottom: 354 / 388 nm). While the geographical patterns are the same, the absolute values differ slightly with the NRT values (left) slightly higher than the 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. The validation of the product is preliminary and ongoing. The results presented here represent the reprocessed 2018 data (OFFL) that were generated after an extensive update of the product. Previously, the computation time of the ALH was so large that only a very small set of a few hundred selected pixels could be processed. This has been resolved by an extensive update of the forward component of the algorithm fit, which now allows global processing of data in near-real time. The reprocessed data have been validated before release, which is what is presented here.

The ALH is still 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.

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 11**, 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 11 – 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]

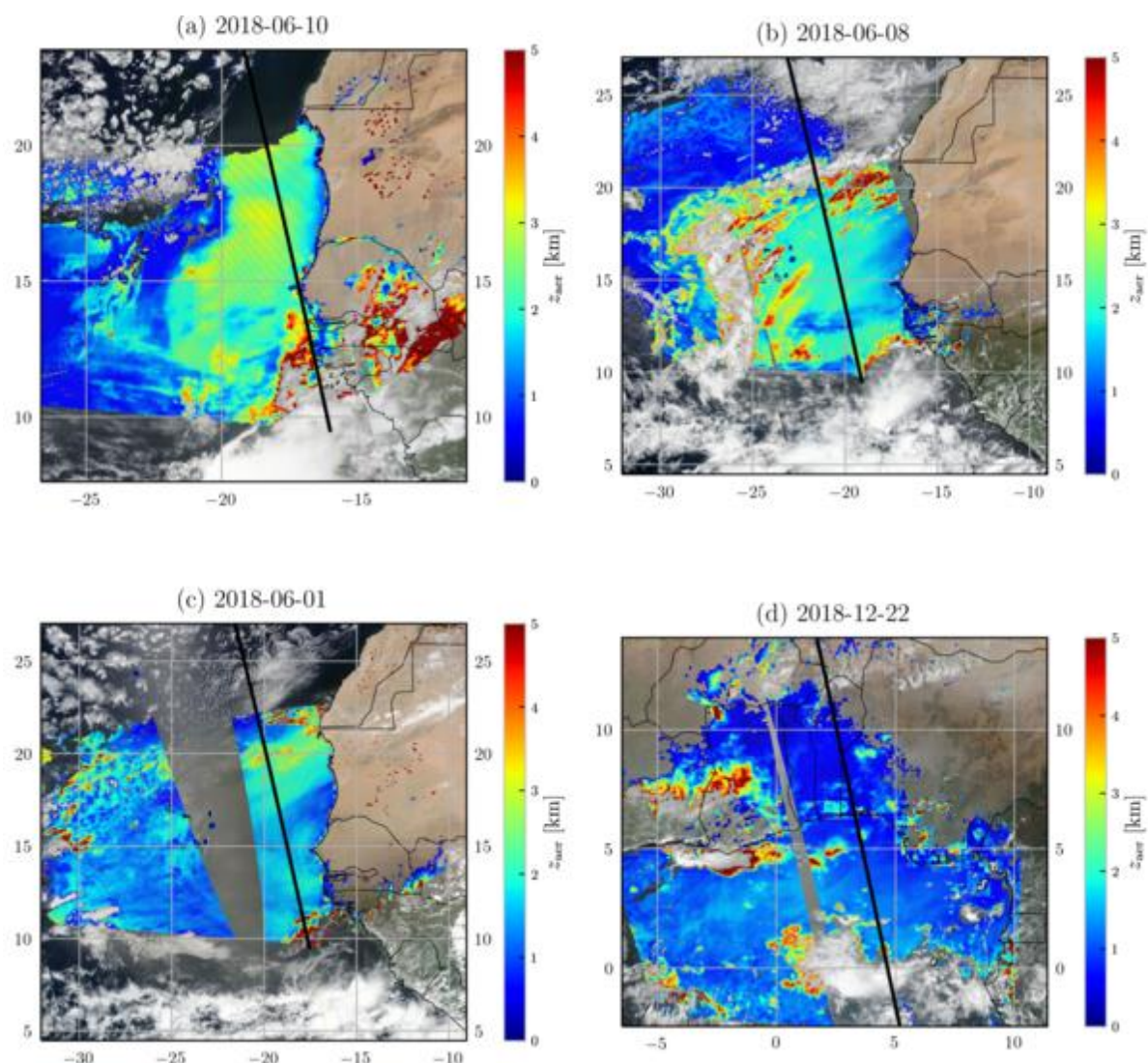


Figure 108: Details of the selected validation cases, showing the ALH on a VIIRS RGB background. The black line represents the Calipso track.

Figure 108 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 Calipso tracks, which have a good coverage of the events. The curtain plots from CALIOP are shown in **Figure 109**, 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 109** 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.

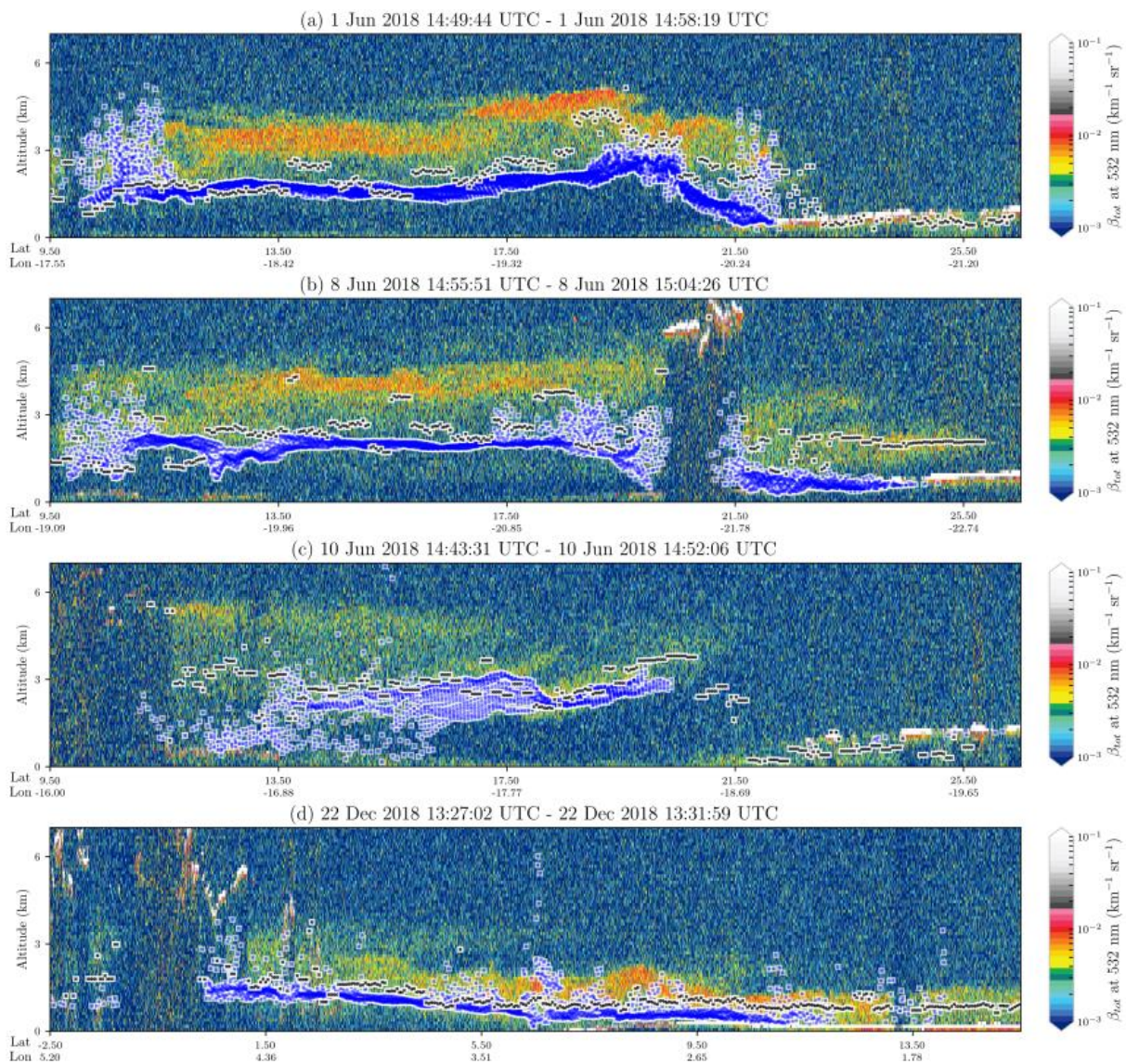


Figure 109: Curtain plots from CALIOP, showing the total attenuated backscatter at 532 nm for the four cases in **Figure 108**. Overplotted on the coloured background image of the total attenuated backscatter are the weighted extinction heights as derived from the backscatter coefficient in black-and-white, and the ALH from collocated TROPOMI pixels in blue-and-white.

In **Figure 110** 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.

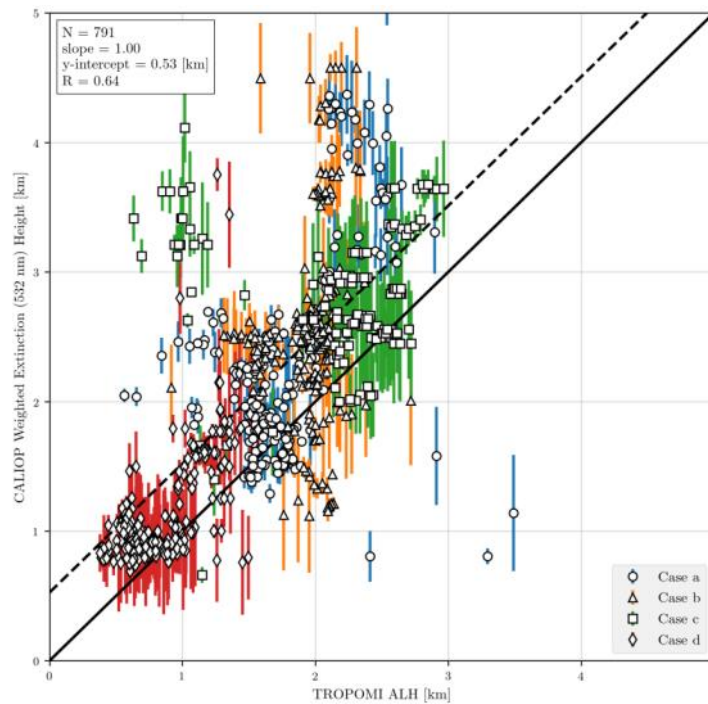


Figure 110: Comparison of ALH from TROPOMI and CALIOP for the cases presented in **Figure 108**. Each case is colour coded. CALIOP weighted extinction height is consistently lower than TROPOMI ALH.

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 111** (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 collocations 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**

111 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.

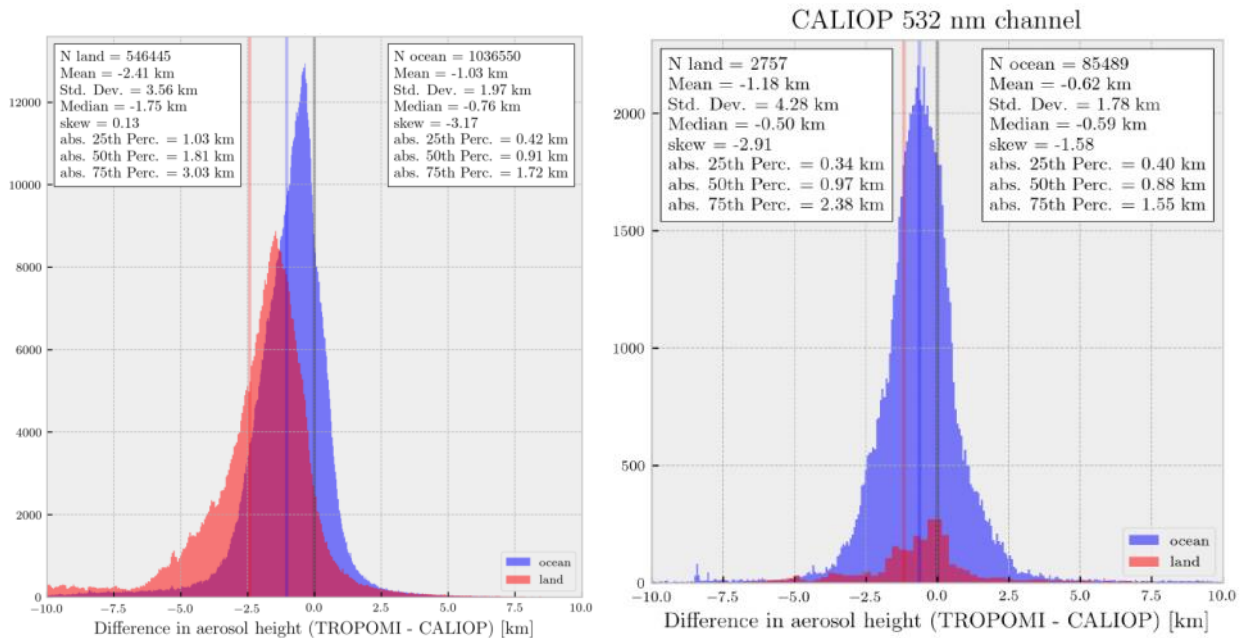


Figure 111: Histogram of differences between CALIOP weighted extinction height and TROPOMI ALH from collocated data between 1 May 2018 and 28 February 2019 (left). The right panel shows the same histogram, but for pixels which were screened for a minimal cost function (chi-squared) smaller than 1E5.

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 112** and **Figure 113**. 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.

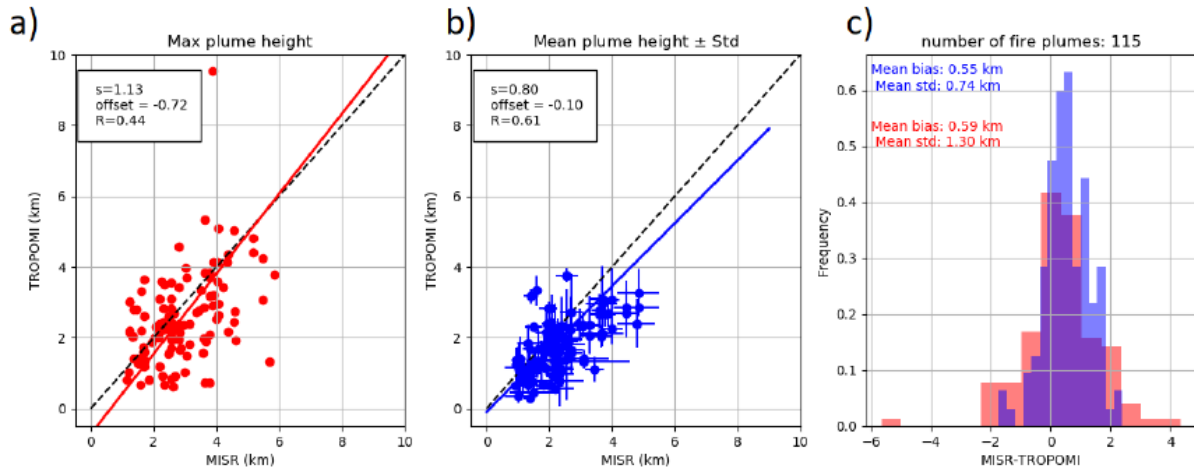


Figure 112: Comparison of TROPOMI ALH and MISR plume height for 115 fires over Northern America in 2018. See Griffin et al, 2019 for details.

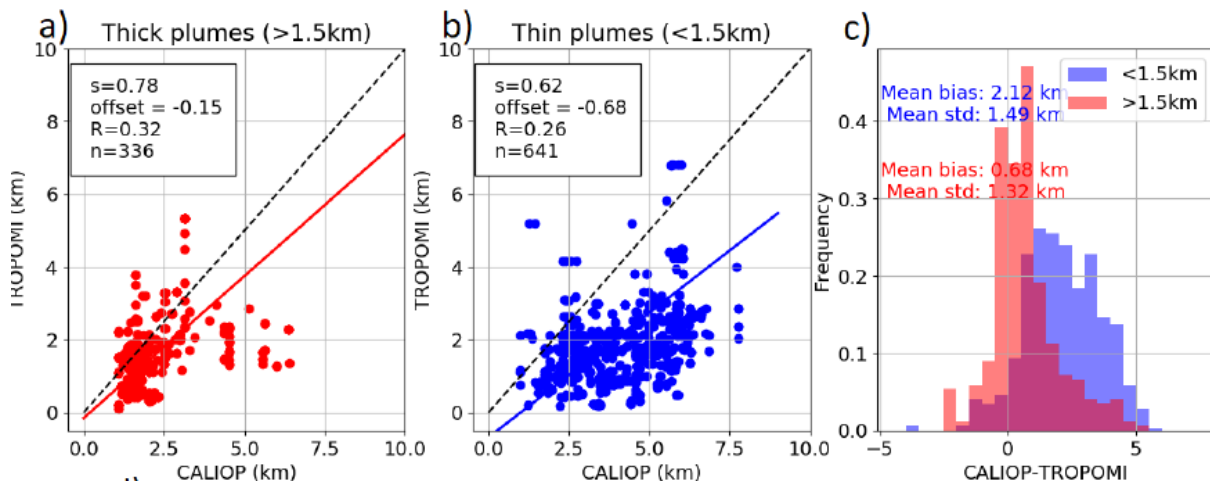


Figure 113: Comparison of TROPOMI ALH and CALIOP average aerosol layer height (top minus bottom of aerosol layer as defined by the feature mask) for collocated pixels near fires over Northern America in 2018. See Griffin et al, 2019, for details.

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 114**). 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.

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.

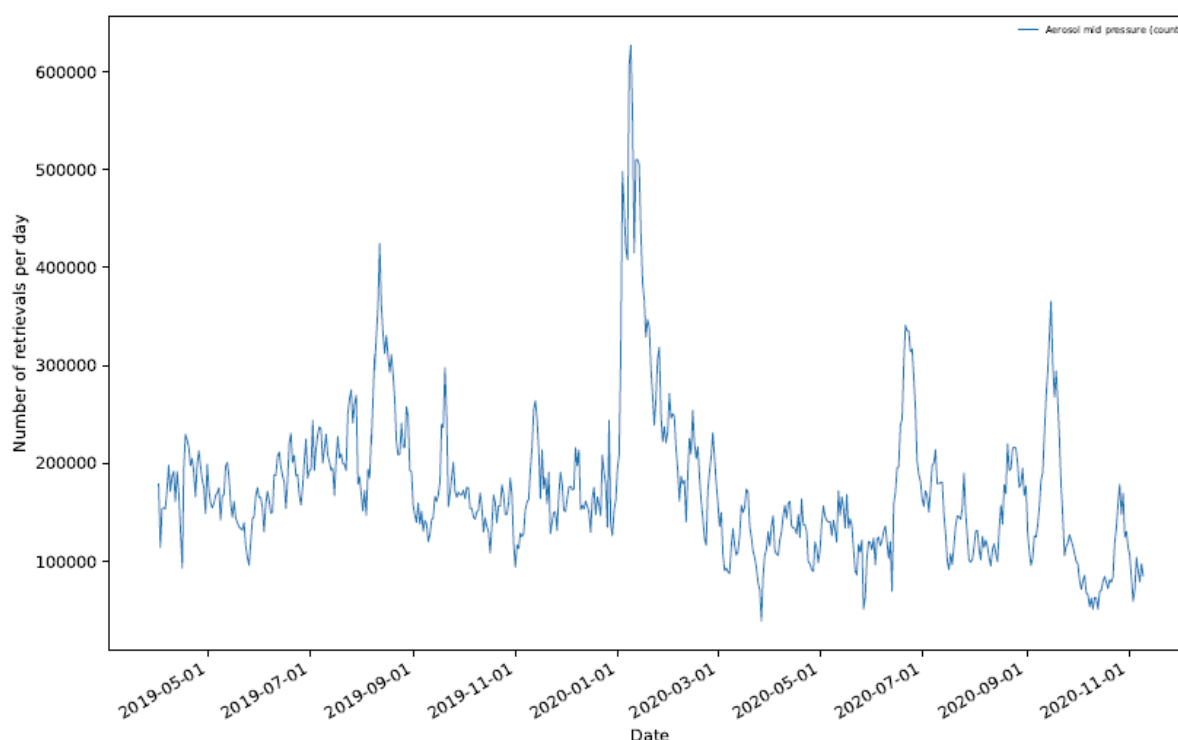


Figure 114: Number of successfully processed ALH pixels per day from April 2019 to November 2020.

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 115** 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.

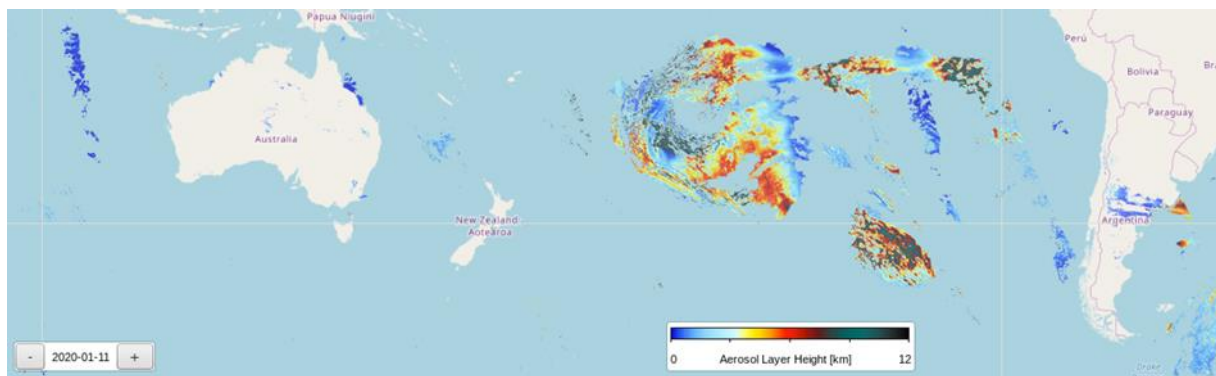


Figure 115: TROPOMI AER_LH product on 11 January 2020 over the south Pacific, showing the altitude as derived by the S5P AER_LH algorithm of the fires smoke from Australian bush fires.

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 116** 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 117** the CALIOP total attenuated backscatter at 532 nm is shown for the yellow track shown in **Figure 116**. The plume can be seen around about 44°S and 110°E at an altitude between about 17 and 21 km, which is much higher than the S5P AER_LH. The AER_LH retrievals from TROPOMI are shown in the curtain plot as black and white dots as before. Clearly, the AER_LH is much lower than the altitude of the smoke plume.

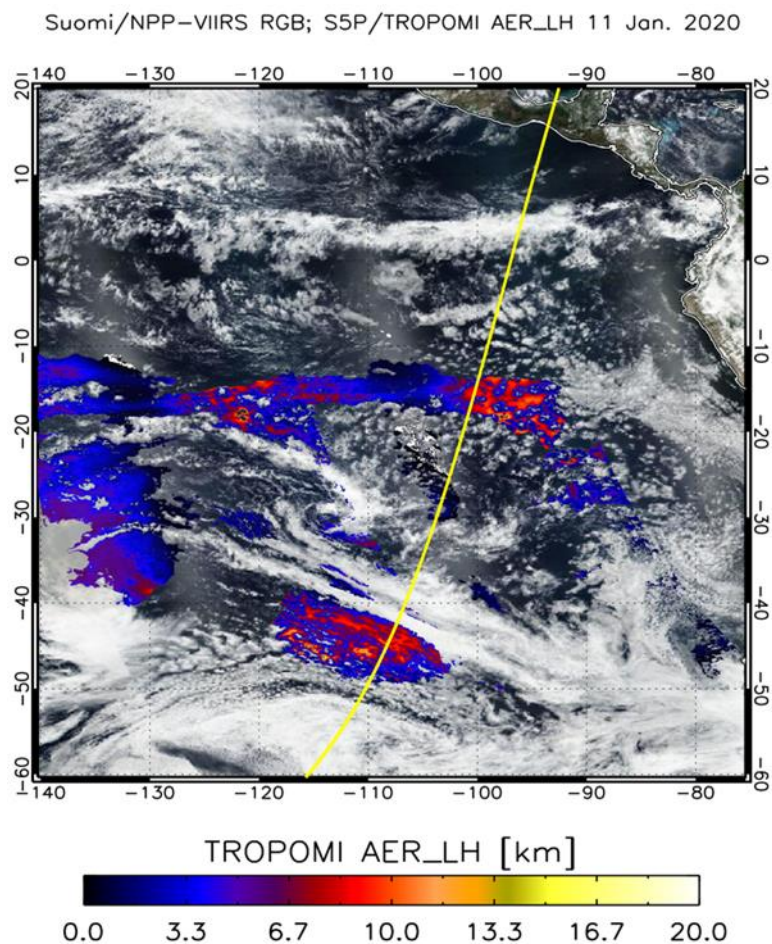


Figure 116: NPP/VIIRS RGB image with S5P/TROPOMI AER_LH on 11 January over the south Pacific with the CALIPSO track of that day overlaid in yellow

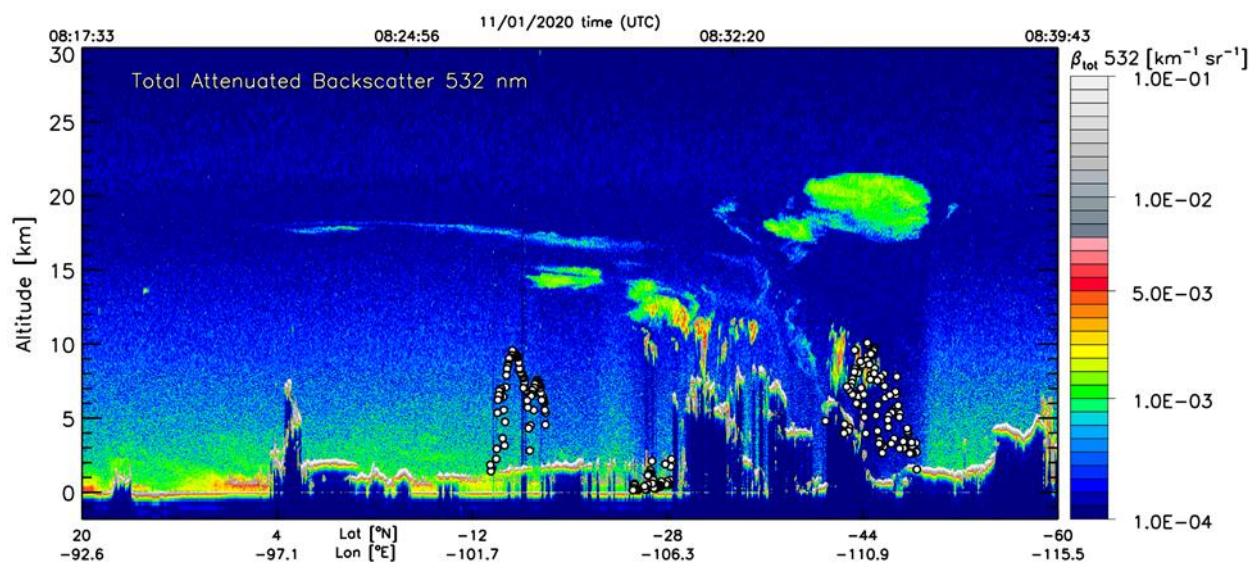


Figure 117: CALIOP aerosol backscatter rat 532 nm along the track shown in **Figure 116** (eastern of Australia on 11 January 11, 2020).

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 116** and **Figure 117**, 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. Compennolle, et al.: Impact of coronavirus outbreak on NO₂ 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.

- Compernelle, 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. Discuss., <https://doi.org/10.5194/amt-2020-122>, in review, 2020.
- Compernelle, 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., PETERS, A., Peters, E., Pommereau, J.-P., Remmers, J., Richter, A., van Geffen, J., Van Roozendaal, M., Wagner, T., and Lambert, J.-C.: Validation of Aura-OMI QA4ECV NO₂ Climate Data Records with ground-based DOAS networks: role of measurement and comparison uncertainties, submitted to Atmos. Chem. Phys., <https://doi.org/10.5194/acp-2019-877>, in review, 2020.
- 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 Roozendaal, 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 Roozendaal, 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 Roozendaal, 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 Roozendaal, 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, NO₂ 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 Gijssel, 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 Compernelle, 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 Veeffkind, 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. Compernelle, 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., Compernelle, 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 NO₂ 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](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 Veeffkind, P.: In-flight calibration results of the TROPOMI payload on-board the Sentinel-5 Precursor satellite, *Atmos. Meas. Tech. Discuss.*, <https://doi.org/10.5194/amt-2019-488>, revised manuscript accepted for publication, 2020.
- Nanda, S., Veeffkind, 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. PETERS, 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 Gijssels, 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., Puertedura, O., Yela, M., Arlander, W., Kåstad Høiskar, B., Tellefsen, C., Karlsen Tørnkvis, 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., Puertedura, O., Yela, M., Braathen, G., Stebel, K., Tørnkvis, 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 OCIO 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. Compernelle, 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. Elokho, 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 NO₂ measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks, *Atmos. Meas. Tech. Disc.*, <https://doi.org/10.5194/amt-2020-119>, in review, 2020.
- 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_2ndS5PVT]	Second S5PVT Meeting and First Results Workshop (including link to presentations)	https://atpi.eventsair.com/QuickEventWebsitePortal/2nd-sentinel-5-precursor-validation-team-and-early-results-meeting/website
[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/utis/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_BEAT]	Basic Envisat Atmospheric Toolbox	http://www.stcorp.nl/beat

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://tcccon-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 12**), 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 12 – 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. Compernelle (BIRA-IASB) H. Eskes (KNMI)
	NO ₂ total column		G. Pinardi (BIRA-IASB) P. Valks (DLR)
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. Compernelle (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. Compernelle (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 and B. Rino 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. Compernelle, 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 Ozone sonde programme (SHADOZ) [ER_SHADOZ], and the Total Carbon Column Observation Network (TCCON) [ER_TCCON], all contributors to WMO's Global Atmosphere Watch (GAW). Data archived in the associated data centres and lists of associated data originators are publicly available.

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

- Rapid delivery O₃ 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. Pazmino, F. Goutail, J.-P. Pommereau, A. Bazureau, and C. Zerefos).
- Rapid delivery O₃ profile data from the SHADOZ network was organised in the framework of the S5PVT AO project CHEOPS-5p (ID #28587, PIs A. Keppens and J.-C. Lambert, BIRA-IASB, Co-Is D. Balis, D. Hubert, W. Steinbrecht, T. Stavrakou, A. Delcloo, S. Godin-Beekmann, T. Leblanc, R. Stübi, A.M. Thompson, T. Verhoelst, G. Ancellet, and V. Duflo). Rapid delivery ozonesonde profile data were also provided by KNMI (A. Piters, M. Allaart) and NOAA (B.J. Johnson).
- Rapid delivery NO₂ data from NDACC MAX-DOAS and ZSL-DOAS stations was gathered in the framework of the S5PVT AO projects CESAR (ID #28596, PI A. Apituley, KNMI) and NIDFORVAL (ID #28607, PI G. Pinardi, BIRA-IASB). The LATMOS SAOZ_RT team (A. Pazmino, A. Bazureau, F. Goutail, and J.-P. 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 NO₂ data from the Pandonia Global Network (PGN).

- Rapid delivery HCHO data from NDACC FTIR and MAX-DOAS stations was gathered in the framework of the S5PVT AO projects CESAR (ID #28596, PI A. Apituley, KNMI) and NIDFORVAL (ID #28607, Co-PIs G. Pinardi and C. Vigouroux, BIRA-IASB). This work could not be possible without the work of the FTIR and DOAS spectra and/or data providers: Carlos Augusto Bauer Aquino (IFRO); Cornelis Becker (SAHO); Thomas Blumenstock and Amelie Röhling (KIT, IMK-ASF); Martine De Mazière, Christian Hermans, François Hendrick, Michel Van Roozendael and Minqiang Zhou (BIRA); Omaira García (AEMET); Michel Grutter, Claudia Rivera, Alejandro Bezanilla, César Guarín 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 Pitters (KNMI); Thomas Wagner, Sebastian Donner and Julia Remmers (MPIC).
- Rapid delivery CO and CH₄ data from TCCON FTIR stations was gathered in the framework of the S5PVT AO project TCCON4S5P (ID #28603, PI M. Kumar Sha, BIRA-IASB).
- Rapid delivery of NDACC data is partly supported by the CAMS-27 data procurement service 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 O₃, NO₂, 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, O₃ 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
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