Sentinel-5 Precursor
Mission Performance Centre

Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products #06: April 2018 – February 2020

Prepared by	S5P Mission Performance Centre
Reference	S5P-MPC-IASB-ROCVR-06.0.1-20200330
CI identification	DI-MPC-ROCVR
Document update	#06
Issue	06.0.1
Date of issue	2020-03-30
Status	Final
Distribution	Public

Approval Record		
Checked by:	J.-C. Lambert (BIRA-IASB) Q. L. Kleipool (KNMI) D. Loyola (DLR) J. P. Veefkind (KNMI) A. Dehn (ESA)	MPC ESL-VAL Lead MPC ESL-L1 Lead MPC ESL-L2 Lead MPC Technical Manager MPC Project Officer
Approved by:	A. Dehn (ESA)	ESA Data Quality Manager
Signature:	J.-C. Lambert (BIRA-IASB)	01-04-20 X Jean-Christopher Lambert Jean-Christopher Lambert Signed by: Jean-Christopher Lambert (Authentication)

Document Information

Document Identification	
Title	Quarterly Validation Report of the Copernicus Sentinel-5 Precursor Operational Data Products #06: April 2018 – February 2020
Type of document	S5P MPC Routine Operations Consolidated Validation Report (ROCVR)
Document ID	S5P-MPC-IASB-ROCVR-06.0.1-20200330
ROCVR update Issue number Date of issue Status	#06 version 06.0.1 2020-03-30 Final
Available at:	http://mpc-vdaf.tropomi.eu
Editors	J.-C. Lambert and A. Keppens (BIRA-IASB)
Contributors	S. Compernolle (BIRA-IASB), K.-U. Eichmann (IUP-UB), M. de Graaf (KNMI), D. Hubert (BIRA-IASB), A. Keppens (BIRA-IASB), Q. Kleipool (KNMI), B. Langerock (BIRA-IASB), M.K. Sha (BIRA-IASB), T. Verhoelst (BIRA-IASB), T. Wagner (MPI-C), C. Ahn (NASA/GSFC), A. Argyrouli (DLR), D. Balis (AUTH), K.L. Chan (DLR), I. De Smedt (BIRA-IASB), H. Eskes (KNMI), A.M. Fjæraa (NILU), K. Garane (AUTH), J.F. Gleason (NASA/GSFC), F. Goutail (LATMOS), J. Granville (BIRA-IASB), P. Hedelt (DLR), K.-P. Heue (DLR), G. Jaross (NASA/GSFC), ML. Koukouli (AUTH), J. Landgraf (SRON), R. Lutz (DLR), S. Nanda (KNMI), S. Niemejer (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), 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 #06: April 2018 – February 2020. Lambert, J.-C., S. Compernolle, K.-U. Eichmann, M. de Graaf, D. Hubert, A. Keppens, Q. Kleipool, B. Langerock, M.K. Sha, T. Verhoelst, T. Wagner, C. Ahn, A. Argyrouli, D. Balis, K.L. Chan, I. De Smedt, H. Eskes, A.M. Fjæraa, K. Garane, J.F. Gleason, F. Goutail, J. Granville, P. Hedelt, K.-P. Heue, G. Jaross, ML. Koukouli, J. Landgraf, R. Lutz, S. Nanda, S. Niemejer, A. Pazmiño, G. Pinardi, J.-P. Pommereau, A. Richter, N. Rozemeijer, M. Sneep, D. Stein Zweers, N. Theys, G. Tilstra, O. Torres, P. Valks, C. Vigouroux, P. Wang, and M. Weber. S5P MPC Routine Operations Consolidated Validation Report series, Issue #06, Version 06.0.1, 154 pp., March 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. This Routine Operations Consolidated Validation Report (ROCVR) integrates results from the MPC Validation Data Analysis Facility (VDAF) consortium [ER_VDAF] with ad hoc support from S5P Validation Team (S5PVT) AO projects [ER_S5PVT]. The MPC 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 to be followed. The present report covers the period of S5P routine operation data from April 2018 till February 2020. It includes validation of Level-2 data quality after the operations switch from the initial ground pixel size of 7 (along track) x 3.5 (across track) km² to the smaller size of 5.5 x 3.5 km², activated on August 6, 2019. It also includes validation results for the S5P L2_AER_LH aerosol layer height data product released to the public since the end of September 2019.

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 ±2.5%. The instrumental switch to smaller (along-track) ground pixels on the 6^{th} of August 2019 does not show any effect on the agreement with the ground-based reference data.

The difference between S5P TROPOMI and other satellite data sets (GOME-2B, OMI, OMPS) over cloudy scenes highlights differences in the cloud models used in the retrieval algorithms. Larger and/or systematic differences between satellite datasets also exist at high solar zenith angles (and 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 +14% or +2.7 DU) and the mean dispersion of the relative difference (about 23% or 4.3 DU) both comply with mission requirements, that is, a bias lower than 25% and an uncertainty (dispersion) less than 25%.

However, during the first year of S5P operations and at several ozonesonde stations, the bias exhibits seasonal patterns with amplitude exceeding mission requirements. During the 2018 biomass burning season the positive bias w.r.t. ozone soundings around the Atlantic basin reaches peak values of up to 10-15 DU (or 40-60%). The onset of the positive bias period is seen again during the 2019 season at Paramaribo. Delivery of ozonesonde data at other Atlantic sites is expected within three months' time at latest. Unphysical biases of 1-2 DU between neighbouring latitude bands occur frequently in the S5P L2_O3_TCL data product. The progression of the orbital sampling by the S5P instrument imprints another spatio-temporal bias pattern that is more elusive and harder to quantify.

Nitrogen Dioxide

The S5P L2_NO2 (NRTI, OFFL, RPRO) data products (tropospheric, stratospheric, and total column) up to version 01.03.02 are in good overall agreement with correlative ground-based measurements from the Pandonia Global Network (PGN), the NDACC ZSL-DOAS/SAOZ, and the MAX-DOAS monitoring networks, and with corresponding satellite data products (OMI). This assessment covers the period of S5P L2_NO2 data from the start of phase E2 (April 2018) until February 2020. In general, a low bias is detected for most of the data products.

The S5P L2_NO2 stratospheric NO_2 column data are compared with the NDACC ZLS-DOAS ground-based measurements at 19 stations from pole to pole. Taking diurnal variation into account, S5P stratospheric NO_2 is generally lower by approximately 0.15 Pmolec/cm^2. This bias of roughly -10% is within the S5P mission requirements (0.2-0.4 Pmolec/cm^2), depending on latitude and season. Also, the dispersion is within mission requirements (0.3 Pmolec/cm^2), taking into account combined random errors and co-location mismatches.

The S5P L2_NO2 tropospheric NO_2 columns are compared to ground-based column data at 17 MAX-DOAS stations and 10 PGN direct sun stations. The comparisons conclude to a negative bias of roughly -30%. This is within the mission requirement of 50%. A dispersion of less than 3 Pmolec/cm^2 exceeds the mission precision requirements of 0.7 Pmolec/cm^2. Furthermore, comparisons of S5P with OMI tropospheric NO_2 data show good agreement with differences of about 0.1 Pmolec/cm^2 (roughly 3%).

The S5P L2_NO2 total NO_2 column data are compared to ground-based Pandora column data at 19 sites of the PGN network. The mean bias is -20% with a station-to-station scatter (dispersion) of 30%.

Ground-based validation show rather similar bias and uncertainty estimates for the L2_NO2 NRTI and L2_NO2 OFFL/RPRO datasets.

Formaldehyde

The S5P L2_HCHO (OFFL, RPRO) formaldehyde tropospheric column product is in good overall agreement to independent ground-based measurements from the NDACC FTIR and MAX-DOAS monitoring networks and to similar satellite data products (OMI, GOME-2B).

The low bias of S5P with respect to NDACC MAX-DOAS measurements is roughly -44% for direct comparisons, and -25% when the S5P and MAX-DOAS averaging kernels are applied mutually. S5P also exhibits a similarly low bias at 25 NDACC FTIR stations with high HCHO columns (-31% for HCHO > 8 Pmolec/cm^2), while a positive bias is observed over clean FTIR sites (+26% for HCHO < 2.5 Pmolec/cm^2). Those bias estimates are within the mission requirements (bias below 80%).

The dispersion of about 9 Pmolec/cm^2 with respect to MAX-DOAS data at 2 sites and of 7 Pmolec/cm^2 with respect to FTIR data at only clean NDACC stations is within the uncertainty mission requirements of 12 Pmolec/cm^2. These values are based on the use of median and median absolute deviation to reduce the influence of large outliers and vertical smoothing.

The S5P bias with respect to OMI is less than -10% (-5% for QA4ECV OMI v12) for most regions, with some larger negative biases in Europe, North America and China (< 30%). The dispersion is below 2 Pmolec/cm^2.

Ground-based validation show similar bias and uncertainty (dispersion) estimates for the L2_HCHO NRTI and L2_HCHO OFFL/RPRO datasets.

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 SO_2 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.

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 +7% (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 -0.27% 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 0.6 is rather low, partly because not all outlying pixels are filtered with the `qa_value` above 0.5.

Clouds

The S5P L2_CLOUD (NRTI and OFFL) cloud height data and cloud top height data compare favourably with ground-based measurements from the CLOUDNET and ARM networks. For about half of the 17 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. Therefore, we consider present validation results as positive.

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. At low cloud fraction, there is a higher discrepancy in cloud height between S5P CLOUD CRB, S5P CLOUD CAL and S5P FRESCO.

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 MODIS on the EOS-Aqua satellite. Furthermore, a direct comparison for multiple days with co-located and re-gridded VIIRS cloud top height and cloud optical thickness has been performed.

Two non-physical geographical patterns in cloud parameters were identified: 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. Tests with the prototype L2_CLOUD processor version 2.x.x (to be operational later in 2020), indicate that in the future these patterns will be 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 will be addressed with the next update in S5P L1B data foreseen in 2020.

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.

Processing Baseline Identification

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

Product ID	Stream	Version	In operation from (orbit #, date)	In operation until (orbit, date)
L1B_RA1/2/…/8		01.00.00	2818, 2018-04-30	current version
L1B_IR_UVN/SIR		01.00.00	2818, 2018-04-30	current version
L2_O3	NRTI	01.00.00 01.01.02 01.01.05 01.01.06 01.01.07	2955, 2018-05-09 4245, 2018-08-08 5931, 2018-12-05 7631, 2019-04-04 7999, 2019-04-30	3943, 2018-07-18 5930, 2018-12-05 7631, 2019-04-04 7999, 2019-04-30 current version
	OFFL	01.01.07	7907, 2019-04-23	current version
	RPRO	01.01.07	2818, 2018-04-30	7906, 2019-04-23
L2_O3_TCL	OFFL (CCD)	01.01.05 01.01.06 01.01.07	2824, 2018-04-30 7435, 2019-03-21 7804, 2019-04-16	7421, 2019-03-20 7791, 2019-04-15 current version
L2_NO2	NRTI	01.00.01 01.00.02 01.01.00 01.02.00 01.02.02 01.03.00 01.03.01 01.03.02	2955, 2018-05-09 3745, 2018-07-04 3947, 2018-07-18 5336, 2018-10-24 5931, 2018-12-05 7519, 2019-03-27 7999, 2019-03-30 9159, 2019-07-20	3364, 2018-06-07 3946, 2018-07-18 5333, 2018-07-24 5929, 2018-12-05 7517, 2019-03-27 7999, 2019-03-30 9158, 2019-07-20 current version
	OFFL	01.02.00 01.02.02 01.03.00 01.03.01 01.03.02	5236, 2018-10-17 5840, 2018-11-29 7425, 2019-03-20 7907, 2019-04-23 8815, 2019-06-26	5832, 2018-11-28 7424, 2019-03-20 7906, 2019-04-23 8814, 2019-06-26 current version
	RPRO	01.02.02	2836, 2018-05-01	5235, 2018-10-17
L2_HCHO	NRTI	01.00.00 01.01.01 01.01.02 01.01.05 01.01.06 01.01.07	2955, 2018-05-09 3947, 2018-07-18 4245, 2018-08-08 5931, 2018-12-05 7636, 2019-04-04 7999, 2019-04-30	3943, 2018-07-18 4244, 2018-08-08 5929, 2018-12-05 7628, 2019-04-04 7999, 2019-04-30 current version
	OFFL	01.00.00 01.01.01 01.01.02 01.01.05 01.01.06 01.01.07	3202, 2018-05-27 3848, 2018-07-10 4147, 2018-07-31 5832, 2018-11-27 7546, 2019-03-29 7907, 2019-04-23	3847, 2018-07-10 4146, 2018-07-31 5831, 2018-11-27 7545, 2019-03-29 7906, 2019-04-23 current version
	RPRO	01.01.02 01.01.05	2818, 2018-04-30 3627, 2018-06-26	4147, 2018-08-01 5832, 2018-11-28
L2_SO2	NRTI	01.00.00 01.01.01 01.01.02 01.01.07	3202, 2018-05-27 3848, 2018-07-10 4147, 2018-07-31 5833, 2018-12-05	3847, 2018-07-10 4146, 2018-07-31 5832, 2018-12-04 current version
	OFFL	01.00.00	3202, 2018-05-27	3847, 2018-07-10

Product ID	Stream	Version	In operation from (orbit #, date)	In operation until (orbit, date)
		01.01.01 01.01.02 01.01.07	3848, 2018-07-10 4147, 2018-07-31 5933, 2018-11-28	4146, 2018-07-31 5932, 2018-11-28 current version
L2_CO	NRTI	01.02.00 01.02.02 01.03.00 01.03.01 01.03.02	5336, 2018-10-24 5931, 2018-12-05 7519, 2019-03-27 7999, 2019-04-30 8906, 2019-07-03	5929, 2018-12-05 7517, 2019-03-27 7999, 2019-04-30 8906, 2019-07-03 current version
	RPRO[2] OFFL	01.02.02 01.03.01 01.03.02 01.02.00 01.02.02 01.03.00 01.03.01 01.03.02	5236, 2018-10-17 2818, 2018-04-30 2463, 2018-04-04 5346, 2018-10-25 5833, 2018-11-28 7425, 2019-03-20 7907, 2019-04-23 8815, 2019-06-26	5346, 2018-10-25 5832, 2018-11-28 2477, 2018-04-05 5832, 2018-11-28 7424, 2019-03-20 7906, 2019-04-23 8814, 2019-06-26 current version
L2_CH4	RPRO[2] OFFL	01.02.02 01.03.01 01.03.02 01.02.02 01.03.00 01.03.01 01.03.02	0657, 2017-11-28 2818, 2018-04-30 2463, 2018-04-04 5833, 2018-11-28 7425, 2019-03-20 7907, 2019-04-23 8812, 2019-06-26	5346, 2018-10-25 5832, 2018-11-28 2477, 2018-04-05 7424, 2019-03-20 7906, 2019-04-23 8814, 2019-06-26 current version
L2_CLOUD	NRTI	01.01.01 01.01.02 01.01.05 01.01.06 01.01.07	3947, 2018-07-18 4243, 2018-08-08 5931, 2018-12-05 7631, 2019-04-04 7999, 2019-04-30	4242, 2018-08-08 5929, 2018-12-05 7631, 2019-04-04 7999, 2019-04-30 current version
	OFFL	01.01.05 01.01.06 01.01.07	5833, 2018-11-28 7547, 2019-03-29 7907, 2019-04-23	7546, 2019-03-28 7906, 2019-04-23 current version
	RPRO	01.01.05 01.01.07	2818, 2018-04-30 2818, 2018-04-30	5832, 2018-11-28 7906, 2019-04-23
L2_AER_AI	NRTI	01.02.00 01.03.02	5336, 2018-10-24 5932, 2018-12-05	5929, 2018-12-04 current version
	OFFL	01.02.00 01.03.02	5336, 2018-10-24 5932, 2018-12-05	5929, 2018-12-04 current version
L2_AER_LH	NRTI	01.03.02		current version
	OFFL	01.03.00 01.03.01 01.03.02	7425, 2019-03-20 7907, 2019-04-23 8815, 2019-06-26	7906, 2019-04-23 8814, 2019-06-26 current version
	RPRO	01.03.01	2818, 2018-04-30	7424, 2019-03-20

Table 1 – S5P TROPOMI data products and processor versions (NRTI near-real-time and OFFL off-line). Note 1: the operational phase (E2) of the S5P TROPOMI mission starts with orbit #2818. Note 2: RPRO 01.03.01 and 01.03.02 have been used to fill gaps in the 01.02.02 RPRO and therefore processor start end dates are not sequential.

Representative Quality Indicators

Based on the validation results reported in this document, representative values of key quality indicators (bias and spread) have been derived for the following S5P operational data products:

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
	OFFL	O_3 column	0.3%	2%	
L2_O3_TCL	OFFL (CCD)	O_3 tropospheric column	+14%	23%	Positive bias over biomass burning. Geographical imprints of sampling-related biases.
L2_NO2	NRTI	NO_2 troposphere NO_2 stratosphere NO_2 total	-22% -9% 0±50%	3.6 Pmol/cm^2 0.4 Pmol/cm^2 -	Total NO_2 bias depending on total column: pristine areas to slight pollution overestimated, high pollution underestimated.
	OFFL RPRO	NO_2 troposphere NO_2 stratosphere NO_2 total	-23% -6% 0±50%	3.6 Pmol/cm^2 0.3 Pmol/cm^2 -	
L2_HCHO	NRTI OFFL RPRO	HCHO low HCHO high	+26% -31%	7 Pmol/cm^2 13 Pmol/cm^2	Positive bias over clean areas (<2 Pmol/cm^2), negative bias over large emissions (> 8 Pmol/cm^2).
L2_SO2	NRTI	SO_2 column	0.2 DU	0.2 DU	Lack of validation sites in areas with high SO_2.
	OFFL	SO_2 column	0.2 DU	0.2 DU	
L2_CO	NRTI	CO column	6.5%	5%	Along orbit stripes. High pollution underestimated. 5% SZA dependence of bias. Outliers in SAA and other sporadic locations not filtered by `qa_value`. Since July 2019 NRTI similar as OFFL.
	OFFL	CO column	6.5%	5%	
L2_CH4	OFFL	CH4 column	-0.27%	0.6%	Along orbit stripes. Underestimation at low albedo. Remaining outliers with `qa_value`>0.5. 1-4% seasonal and SZA dependence of bias.
L2_CLOUD	NRTI	CAL CTH CRB CH CAL COT	-20% -30% +7.9 [-]	2 km 1 km -	Bias towards the a priori cloud height up to and including 01.01.07. Snow/ice degrades retrievals. Occurrence of C(T)H equal to surface height at low cloud fraction. Across track CTH and CF pattern. North-South cloud albedo pattern. COT positive bias vs VIIRS.
	OFFL	CAL CTH CRB CH CAL COT	-20% -30% +7.9 [-]	2 km 1 km -	
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) as estimated from the validation studies of the S5P TROPOMI operational data products identified in the **Table 1**. The processor version number is not mentioned as the estimates are representative for all versions available publicly. CTH: cloud-top-height; CH: cloud height; COT: cloud optical thickness.

Table of Contents

1 Introduction

1.1 Background information on Sentinel-5 Precursor TROPOMI

TROPOspheric Monitoring Instrument (TROPOMI) [ER_TROPOMI] is the unique payload of the ESA/Copernicus Sentinel-5 Precursor mission (S5P) launched on October 13, 2017. The prime function of TROPOMI is to monitor the global distribution of atmospheric trace gases and aerosols for a better understanding of air quality, the ozone layer, atmospheric chemistry and transport, ultraviolet radiation, and climate change. The instrument is a nadir-viewing hyperspectral spectrometer measuring, in the ultraviolet-visible (270-495 nm), near-infrared (675-775 nm) and shortwave infrared (2305-2385 nm), the solar radiation scattered by the Earth's atmosphere and reflected by the Earth's surface and by clouds, as well as solar spectral irradiance. Daily coverage at the high horizontal resolution of 7 x 3.5 km^2 before and 3.5 x 3.5 km^2 after the operations switch to smaller ground pixel size activated on the 6^{th} of August 2019, is accomplished thanks to a Sun-synchronous polar orbit (equator crossing time of 13:30 local solar time) and a wide swath width of 2600 km across track. From the TROPOMI radiometric measurements of Earth's radiance and solar irradiance, on-ground data processors retrieve the atmospheric abundance of ozone (O_3), nitrogen dioxide (NO_2), formaldehyde (HCHO), sulphur dioxide (SO_2), carbon monoxide (CO), methane (CH_4), as well as cloud and aerosol properties.

The S5P mission is a key component of the space segment of the European Earth Observation programme Copernicus [ER_CoperESA]. As such, it has an operational and service-oriented vocation. With a 7-year operation lifetime, the S5P mission aims at filling in the anticipated observational gap of key atmospheric composition data between, from one part, Envisat SCIAMACHY (operational in 2002-2012), EOS-Aura OMI (operational since 2004) and the EUMETSAT EPS MetOp GOME-2 series (initiated in 2006, with the latest MetOp-C launched in November 2018), and from the other part, the upcoming series of Copernicus Sentinel-4 and Sentinel-5 missions scheduled after 2022.

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 #06 of the MPC ROCVR presents quality information for the S5P operational data products obtained in nominal mode from April 2018 until February 2020. It is structured as follows:

2 S5P Data Quality Requirements

Validation results can be interpreted to evaluate whether or not S5P Level 2 data products meet user requirements. Targets for key quality indicators of the S5P Level 2 data products have been formulated in the *S5P Geophysical Validation Requirements* document ([S5PVT-Req], Page 19) and the *S5P Cal/Val Plan for the Operational Phase* ([S5P-CSCOP], Page 14). 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 dispersion (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: Dispersion
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 approaches, their respective validation is reported in separate subsections.

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 (NRT vs. OFFL).

4.2.3 Field campaigns and modelling support

After a period of passive monitoring, L2_O3 NRTI is now assimilated in the Copernicus Atmosphere Monitoring Service system (CAMS), which also assimilates data from a variety of other ozone column measuring satellite instruments. See Inness et al. (2019) for further details..

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, but nevertheless, caution remains required for qa_value below 0.5. An alternative set of filter criteria for 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.

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) and at the *3rd S5PVT workshop* (ESA/ESRIN, November 11-14, 2019). Individual contributions to the workshops are archived in https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5p-first-product-release-workshop/sentinel-5p and in https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5-precursor-workshop-2019/sentinel-5p, while up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal at http://mpc-vdaf.tropomi.eu. This report also includes quality information provided by the CAMS team at ECMWF.

Current conclusions are valid for the S5P data obtained in the operational phase E2 of the mission, from May 2018 until February 2020, and on the reference data available at the time of this report: typically, until end of December 2019 for the Dobson and Brewer data, and up to beginning of February 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 100 to 9000 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**).

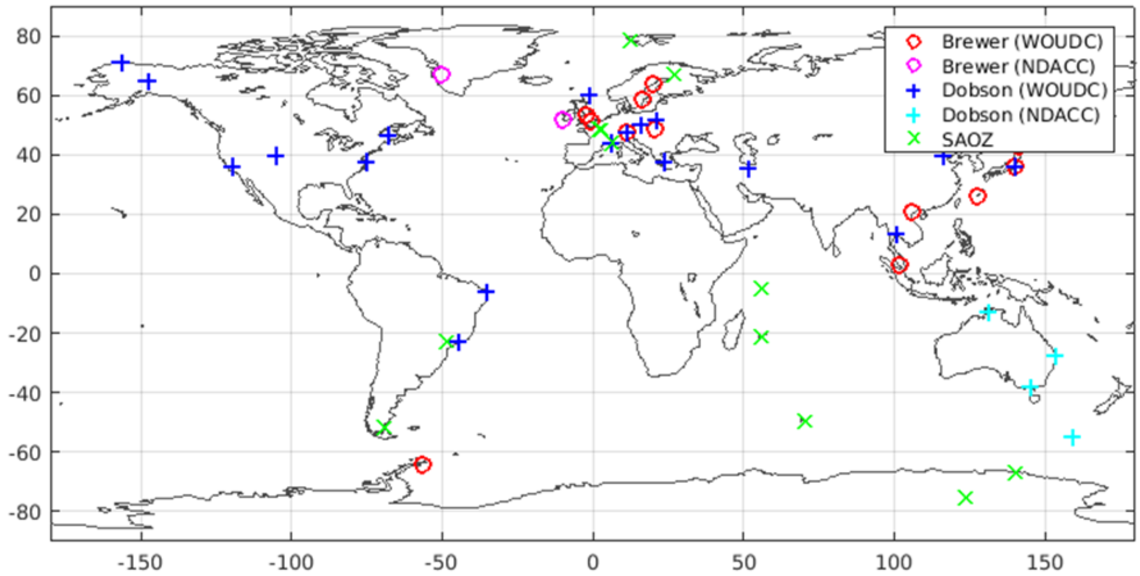


Figure 1: Geographical distribution of Brewer, Dobson and ZSL-DOAS ground-based stations for which suitable co-locations with S5P L2_O3 NRTI ozone data have been identified (May 2018 until February 2020).

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. This median bias value falls well within the mission requirements (max. bias 3.5-5%).

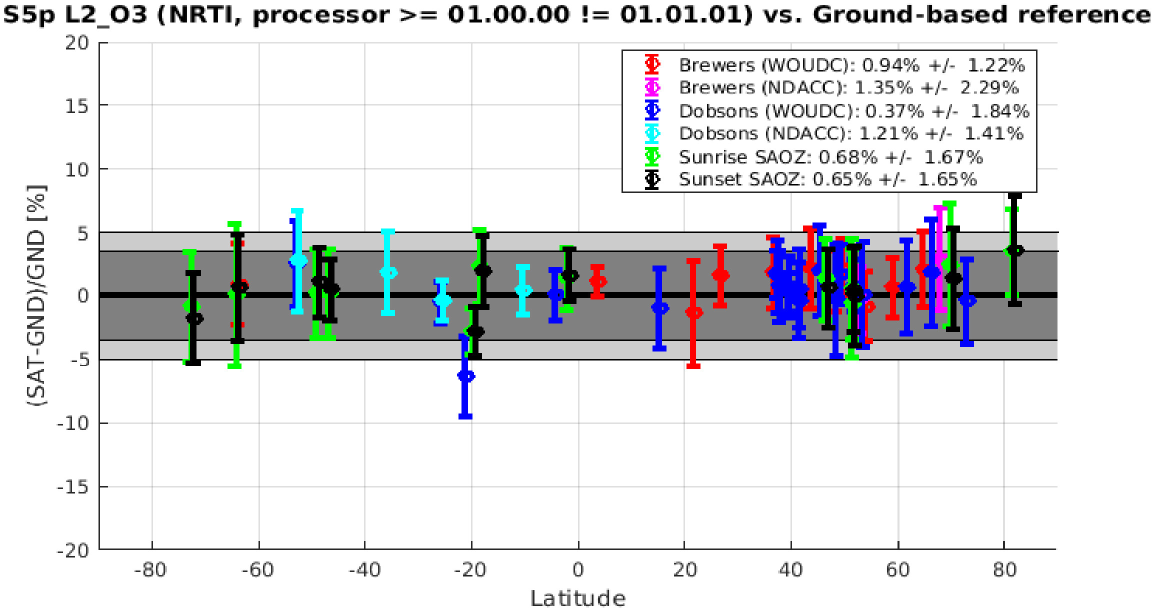


Figure 2: Meridian dependence of the median (the circular markers) and spread (±1 sigma, the error bars) of the percent relative difference between S5P TROPOMI L2_O3 (PDGS NRTI processor v1.0.0 up to v1.1.7, 20 February 2020) and ground-based (GND) ozone column data, represented at individual stations 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).

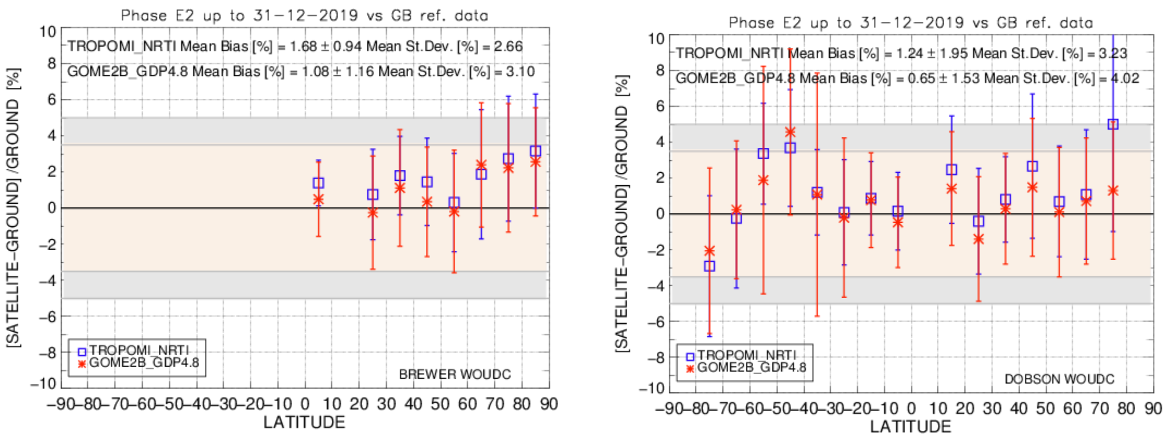


Figure 3: Bias between, from one hand, S5P TROPOMI L2_O3 NRTI (blue) and GOME-2B GDP 4.8 (red), and form the other hand, Brewer (left-hand panel) and Dobson (right-hand panel) network total ozone data (datasets from WOUDC only). The time period of data used for these plots is May 2018 – December 2019 (most recent availability of ground-based data).

4.3.4 Dispersion

The ±1σ dispersion of the difference (between S5P and reference ground-based network data) around their median value rarely exceeds 3-4% for the comparisons with direct-sun instruments (cf. the error bars depicted in **Figure 2**). Combining random errors in both satellite and reference measurements with irreducible co-location mismatch effects, it is concluded that the random uncertainty on the S5P measurements falls within the mission requirements of max.±2.5%.

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°), Air Mass Factor (AMF) and cloud fraction (CF) of the TROPOMI measurement does not reveal any variation of the bias larger than 2% over the range of those influence quantities (**Figure 4**).

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

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

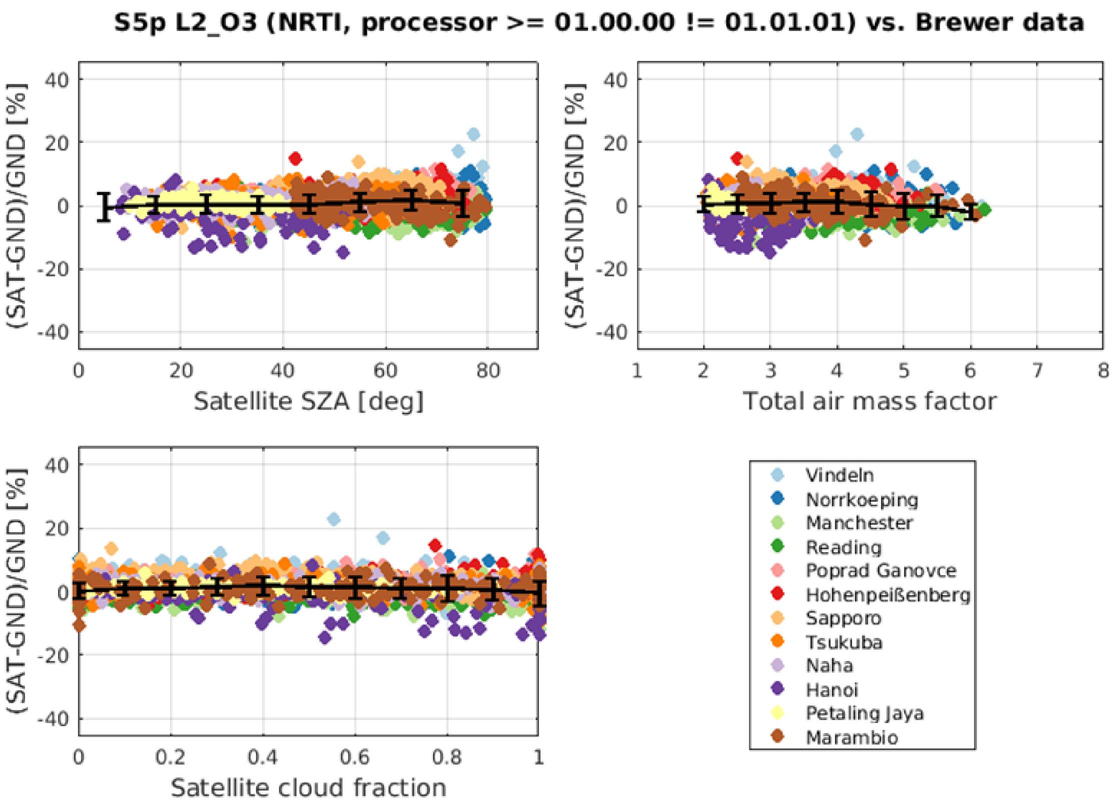


Figure 4: Dependence of the difference between S5P and ground-based Brewer ozone data on the satellite solar zenith angle (SZA), the satellite fractional cloud cover, and the satellite air mass factor, including a mean and standard deviation per 10-degree, 0.1 CF-increment or 1 AMF-increment bin.

4.3.6 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 5**. The overall good agreement is corroborated by Pearson correlation coefficients always above 0.95.

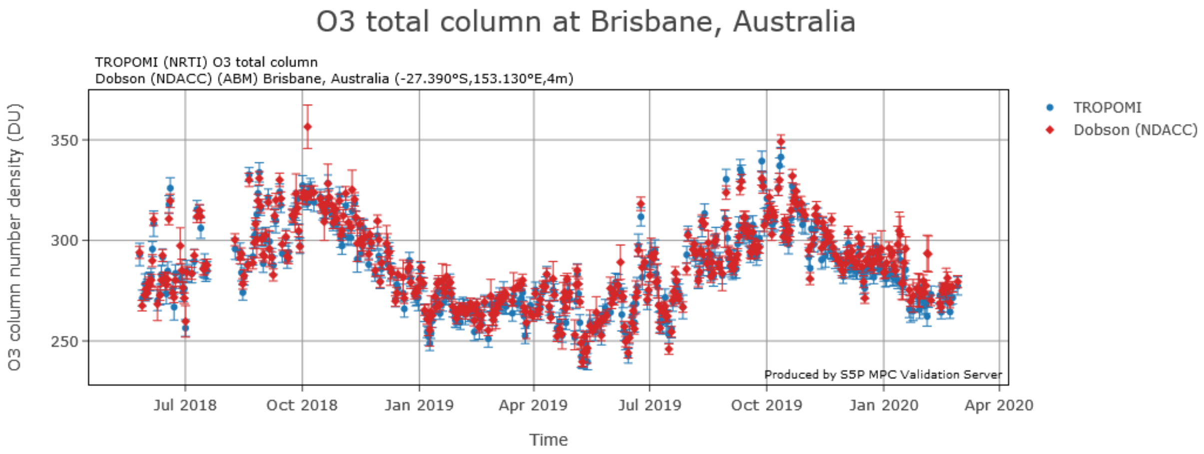


Figure 5: Time series of S5P TROPOMI NRTI and Dobson total ozone data at the NDACC station of Brisbane in Australia (data courtesy M. Tully, ABM).

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 6** where NRTI and OFFL data are compared.

Geographical patterns in the L2_O3 ozone data products – revealed by comparisons with other satellite datasets – are likely to be associated with differences in the processing of the cloud properties, in the use of either a surface albedo climatology or a fitted effective albedo, and, in the case of a comparison of data from two different satellites, to differences in overpass times (3.5 hours difference between S5P and GOME-2).

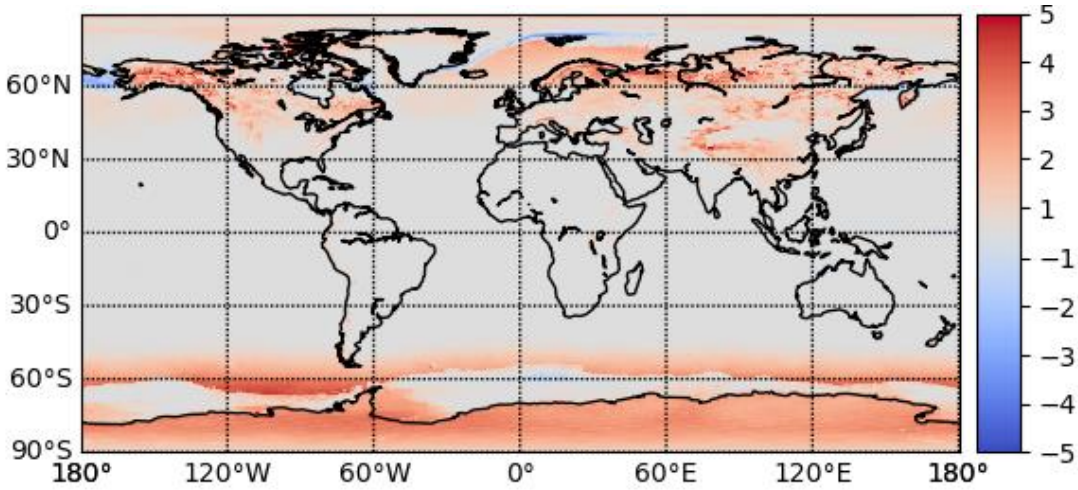


Figure 6: Percent relative difference between S5PL2_O3 total ozone data retrieved with the NRTI and OFFL processors (period November 2017 through September 2018). Geographical patterns in this case cannot be associated with real ozone features but reveal rather the effect of using either a surface albedo climatology (NRTI product) or fitting an effective albedo (OFFL product).

4.3.8 Other features

None to report.

4.3.9 Switch to smaller ground pixel size

On 6 August 2019, the nominal ground pixel resolution of the TROPOMI measurement was reduced to 5.5 x 3.5 km^2, i.e. shorter by 1.5 km in the along-track direction, by reducing the integration time. **Figure 7** shows no evidence of a negative effect on the agreement between satellite and ground-based reference data after this switch in resolution.

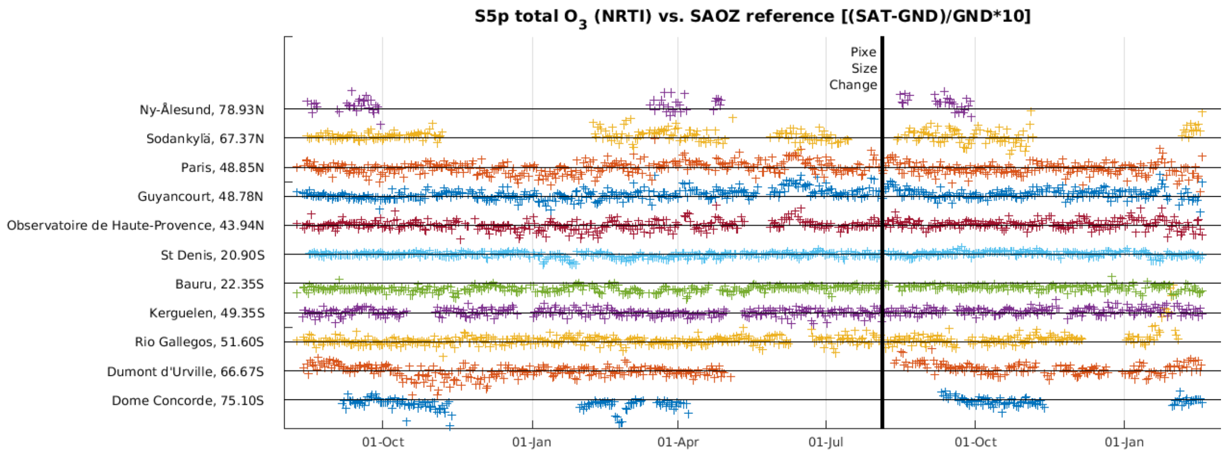


Figure 7: Time series of the difference between the S5P NRTI L2_O3 ozone column data and the SAOZ reference data at 12 stations from pole to pole, from August 2018 until the end of February 2020. The moment of the switch in TROPOMI pixel size (along track reduction from 7 to 5.5 km) is identified by the bold vertical line.

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, 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.

According to validation results this criterion might be relaxed, but nevertheless, caution remains required for qa_value below 0.5. Additional 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.

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) and at the *3rd S5PVT Workshop* (ESA/ESRIN, November 11-14, 2019). Individual contributions to the workshops are archived in https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5p-first-product-release-workshop/sentinel-5p and in https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5-precursor-workshop-2019/sentinel-5p, while up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal at http://mpc-vdaf.tropomi.eu.

Current conclusions are valid for the S5P data obtained in the operational phase E2 of the mission, from May 2018 until February 2020, and on the reference data available at the time of this report: typically, until end of December 2019 for the Dobson and Brewer data, and up to beginning of February 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 100 to 9000 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 8**).

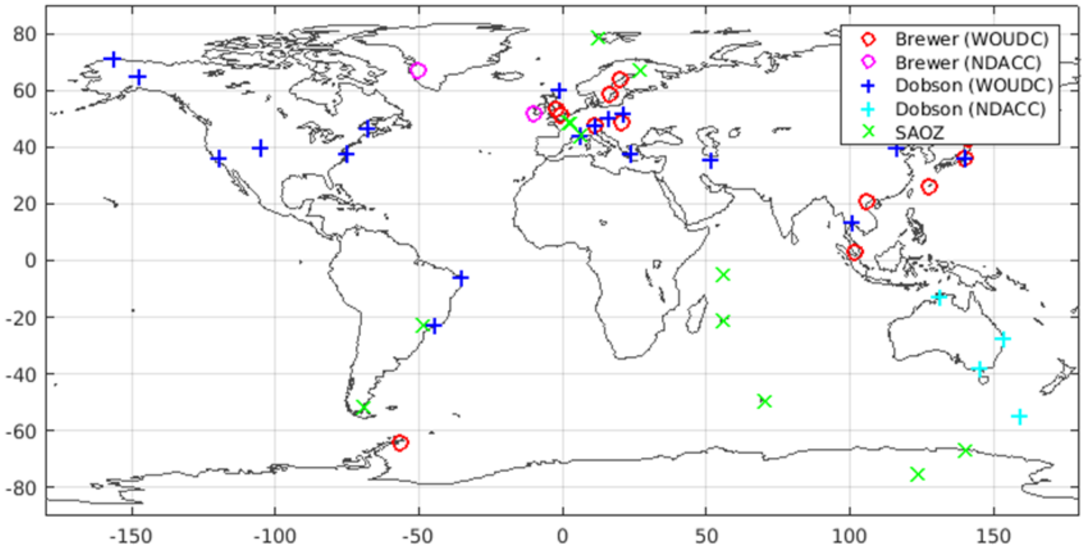


Figure 8: Geographical distribution of Brewer, Dobson and ZSL-DOAS ground-based stations for which suitable co-locations with S5P L2_O3 OFFL ozone data have been identified (period up to February 2020).

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 9**. 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 10**). This median bias value falls well within the mission requirements (max. bias 3.5-5%).

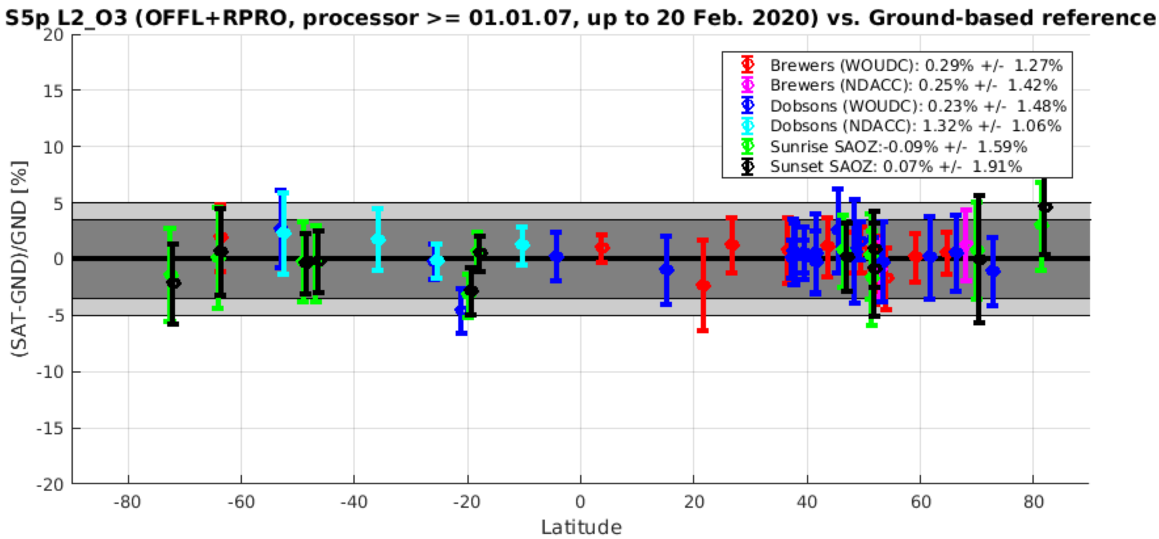


Figure 9: Meridian dependence of the median (the circular markers) and spread (±1 sigma, the error bars) of the percent relative difference between S5P TROPOMI L2_O3 (PDGS OFFL processor v1.1.7) and ground-based (GND) ozone column data, represented at individual stations from the Antarctic to the Arctic 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).

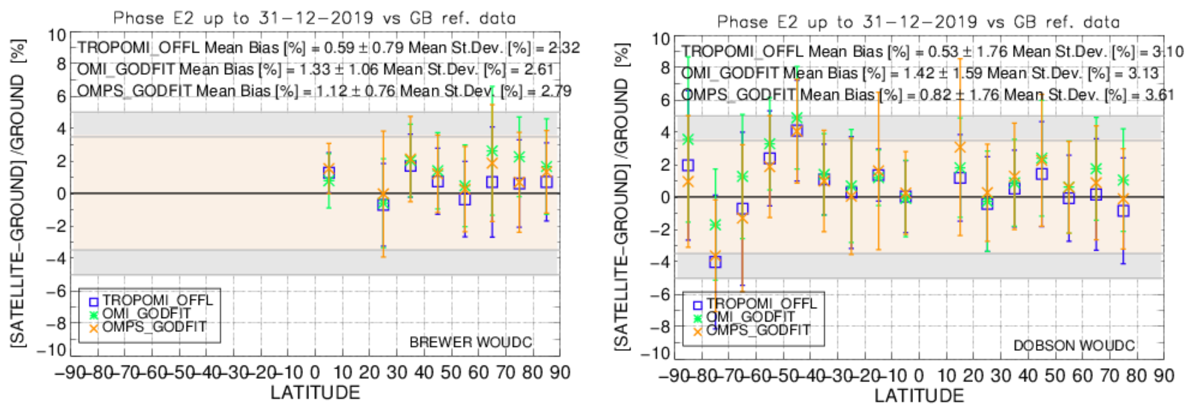


Figure 10: Bias between, on the one hand, S5P L2_O3 OFFL (blue), OMI GODFIT v4 (green) and OMPS GODFIT v4 (red), and on the other hand, the Brewer (left-hand panel) and Dobson (right-hand panel) network total ozone data (datasets from WOUDC only). The time period of data used for these plots is May 2018 – December 2019 (most recent availability of ground-based data).

4.4.4 Dispersion

The ±1σ dispersion of the difference (between S5P and reference ground-based network data) around their median value rarely exceeds 3-4% for the comparisons with direct-sun instruments (cf. the error bars depicted in **Figure 9**). 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%.

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 11**).

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

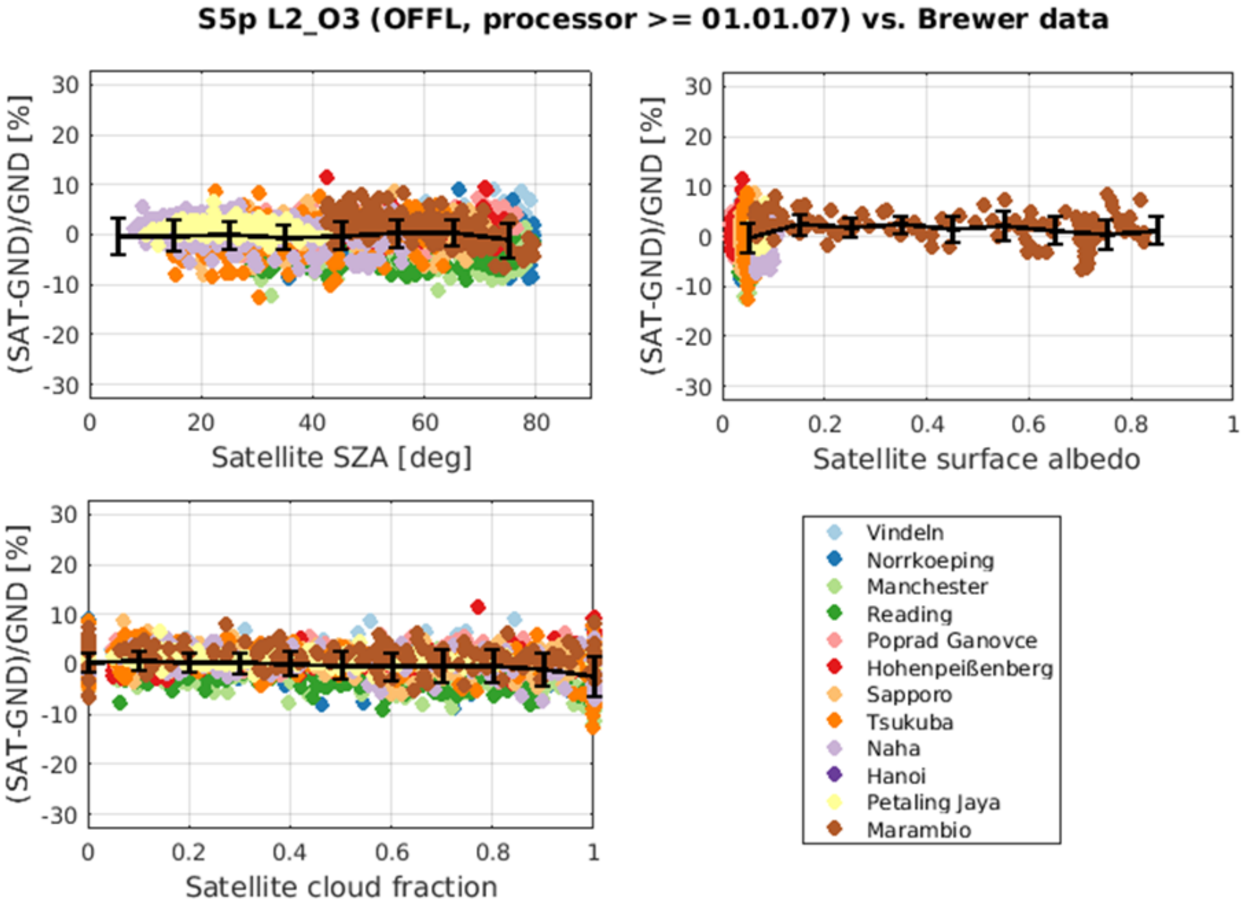


Figure 11: Dependence of the difference between S5P and ground-based Brewer ozone data on the satellite solar zenith angle (SZA), on the satellite fractional cloud cover (CF), and on the surface albedo, including a mean and standard deviation per 10-degree, 0.1 CF-increment and 0.1 albedo increment bin.

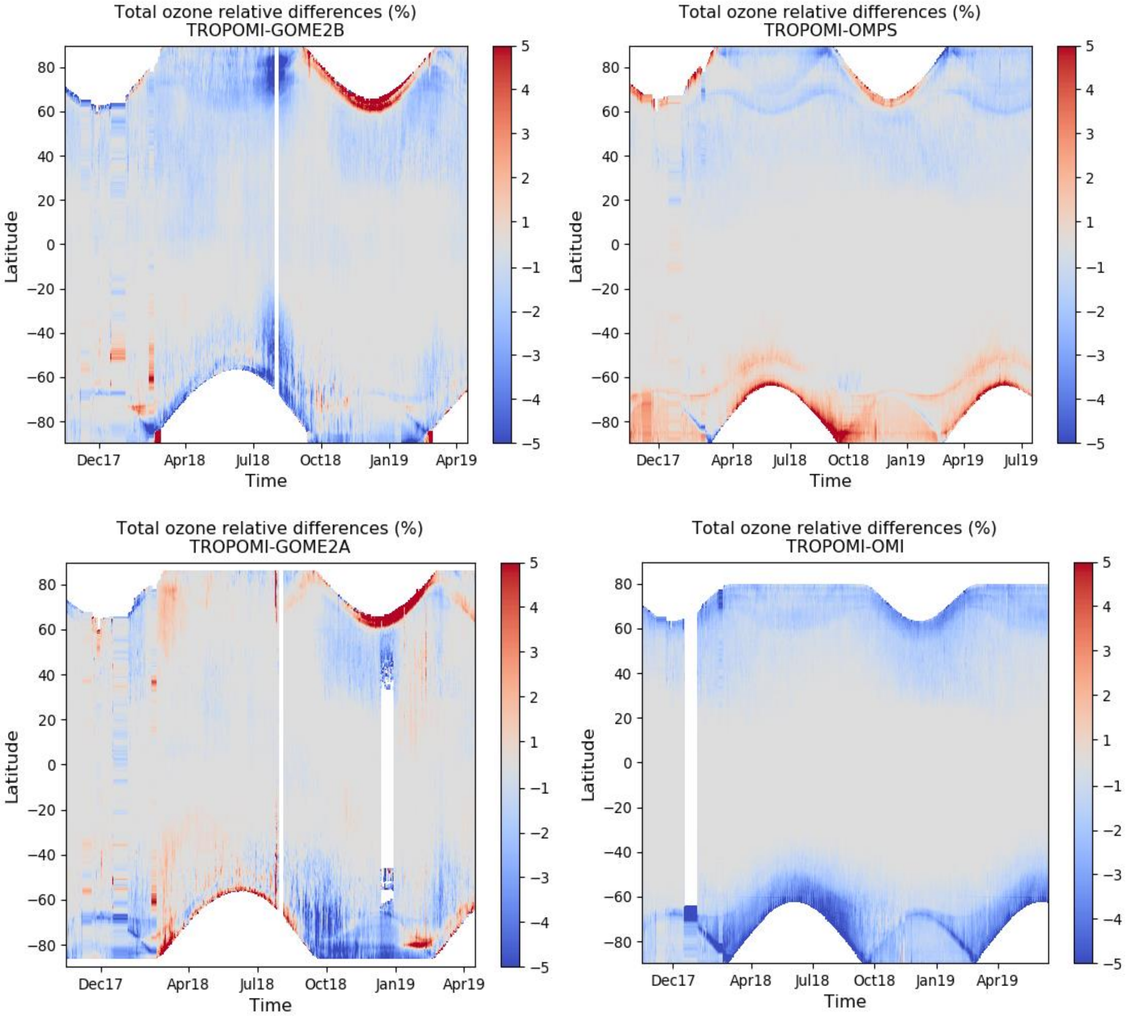


Figure 12: Latitude and time cross-section of the percent relative difference between S5P L2_O3 OFFL total ozone data and comparable total ozone data sets from other satellite: GOME-2B GDP 4.8, S-NPP OMPS, GOME-2A GDP 4.8, and Aura OMI.

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

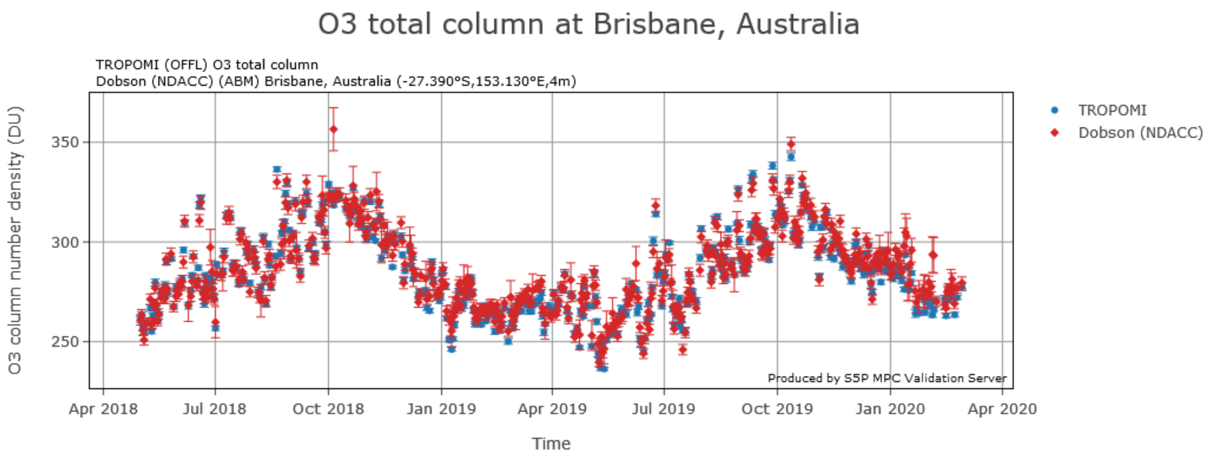


Figure 13: Time series of S5P TROPOMI OFFL and Dobson total ozone data at the NDACC station of Brisbane in Australia (data courtesy M. Tully, ABM).

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 6** where NRTI and OFFL data are compared.

Geographical patterns in the L2_O3 ozone column data products – revealed by comparisons with other satellite datasets – are likely to be associated with differences in the processing of the cloud properties, in the use of either a surface albedo climatology or a fitted effective albedo, and, in the case of a comparison of data from two different satellites, to differences in overpass times (3.5 hours difference between S5P and GOME-2). Moreover, **Figure 12** 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. The SHADOZ data version used here is V06. The ozonesonde profile data are first quality controlled (Hubert et al., 2016) 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 16** and **Figure 18**). 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 – 12 February 2020, the ground-based validation analysis considers 657 S5P OFFL CCD data products and 544 ozonesonde flights at nine sites across the tropics (**Figure 14**). S5P data averaged over the entire tropical region are also intercompared (**Figure 17**) to GOME-2B data (May 2018 – November 2019) and to OMI data (May 2018 – December 2019).

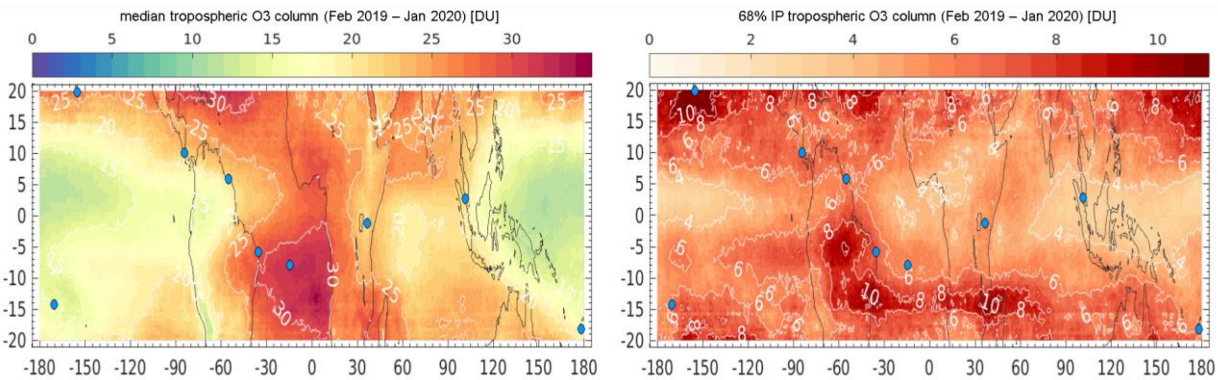


Figure 14: Median value (left) and 68% interpercentile (right) of S5P OFFL tropospheric ozone column data (CCD) over a full year of operations (February 2019 – January 2020). Blue markers locate the nine ground-based ozonesonde stations used in the validation analysis. These maps provide context to **Figure 15** and **Figure 16**.

5.3.3 Bias

S5P tropospheric O_3 column values are on average larger than the ozonesonde values at 8 out of the 9 sites (**Figure 16** and **Figure 18**). The mean bias over the network is +14% or +2.7 DU (**Figure 18**, 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 17**. The agreement with OMI is good, with a mean difference of -0.3 DU or -1%. The larger mean difference of +2.2 DU or +10% compared to GOME-2B might indicate a slight general overestimation of TROPOMI and may also -at least partly- be attributed to the different overpass times of MetOp-B (9:30 desc.) and S5P (13:30 asc.) 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 12-34% or 2.8-7.0 DU (**Figure 16** and **Figure 18**), and the network average is 23% or 4.2 DU (**Figure 18**, centre and bottom right). Dispersion values at five sites are not compliant with the mission requirement for the random component of the uncertainty (<25%). However, all of these sites are located in an area with large natural percentage variability in the tropospheric O_3 field and there is a considerable difference in spatio-temporal sampling between S5P and ozonesonde. In addition, the random component of the uncertainty of the ozonesonde measurement contributes about 5-10% to the observed spread in the differences. Hence, the uncertainty of the S5P data is better than the 23% observed spread in the comparisons to ozonesonde and therefore overall compliant with the mission requirement.

Satellite intercomparisons exhibit a dispersion of 4.2-4.3 DU or ~21% when averaged over the entire tropical belt (**Figure 17**), which is coherent with the average dispersion found in comparisons to the ground-based network.

5.3.5 Dependence on influence quantities

Nothing to report.

5.3.6 Short term variability

S5P time series in **Figure 15** show signs of the passage of biomass burning affected air (high tropospheric O_3 values) at Atlantic and African sites (Heredia, Paramaribo, Natal, Ascension Island, Nairobi) and signs of intense convective activity (low tropospheric O_3 values) at 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 about 25% or 5 DU. This finding is possibly related to a S5P data quality issue and is being monitored during the 2019 biomass burning season. Comparisons to Paramaribo data during July-November 2019 show another period of elevated positive bias. Ozonesonde data at the other Atlantic and African sites are not yet available and are expected by mid-2020. Co-located S5P and reference measurements correlate fairly well for sites with well-sampled comparison time series. Pearson’s skipped correlation coefficients range between 50% and 71% at individual stations, while the network average is 59% (**Figure 18**, top left).

5.3.7 Geographical patterns

Annual median TROPOMI data (February 2019 – January 2020, **Figure 14**) 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 intrusion of ozone-rich air from higher latitudes. It shows the ability of S5P to observe the expected large-scale geographical patterns. At smaller scales, however, two artificial sampling-related patterns are noted.

The CCD algorithm requires an ample sampling of input total O_3 column data to allow a robust estimate of a reference stratospheric O_3 column. This requirement is not always fulfilled. As a result, biases of 1-2 DU between neighbouring latitude bands are found in many S5P data products. The orbital sampling by the S5P instrument progression imprints another, somewhat more elusive spatio-temporal bias pattern that is harder to quantify.

5.3.8 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 Samoa in **Figure 15** 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.

Due to the limited number of ozonesonde comparisons since 6 August 2019 it is currently not possible to have a clear view on the impact of the change in S5P ground pixel size on the quality of S5P tropospheric ozone products. The available comparisons (at 4 stations) show no clear changes (**Figure 14** and **Figure 15**).

A bug in the processor caused incorrect orbit numbers in the S5P product filename for orbits 9918-10387. As a result, 34 L2_O3_TCL products (12 Sep – 15 Oct 2019) were not disseminated to users on the S5P pre-operations data hub. Normal dissemination operations resumed from orbit 10401 onwards (16 Oct 2019).

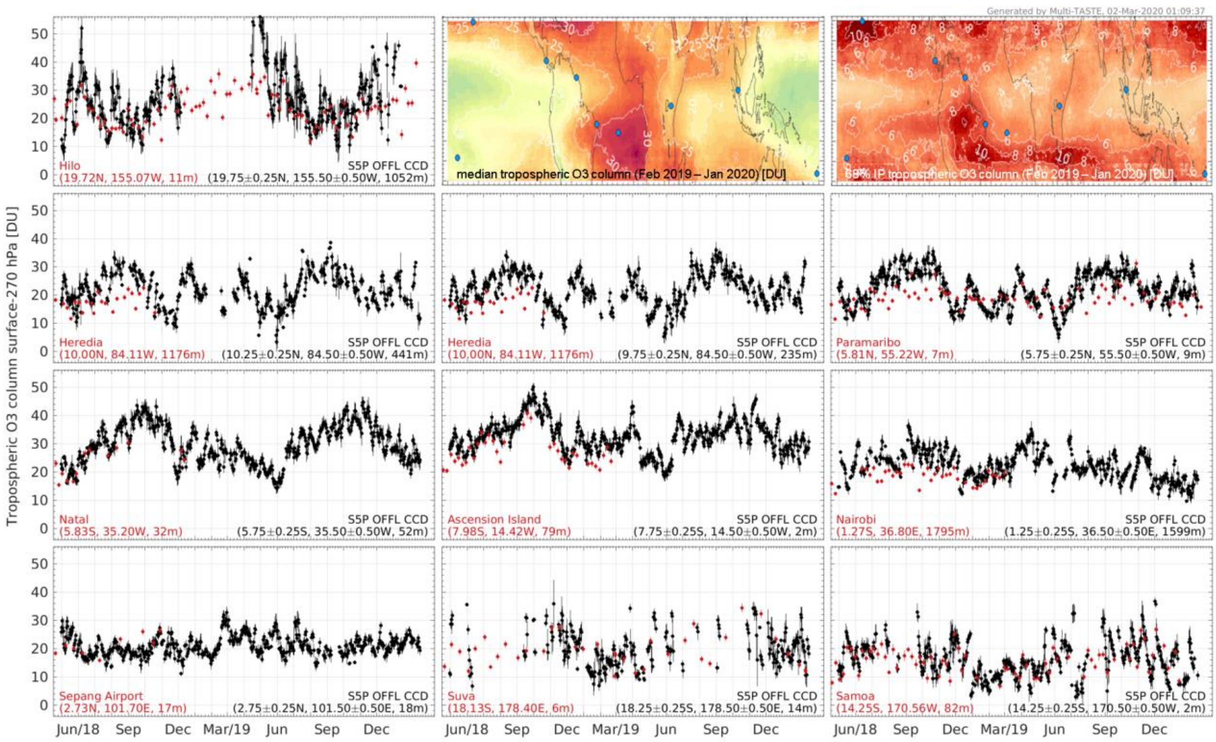


Figure 15: Time series of spatially co-located tropospheric O_3 column data by ozonesonde (red) and by S5P OFFL v01.01.05+v01.01.06+v01.01.07 (black). All data have been screened following recommendations by the data providers. The first row also maps the location of the ozonesonde sites and the characteristics of the tropospheric O_3 field (median and 68% interpercentile over one year of S5P data, see **Figure 14**).

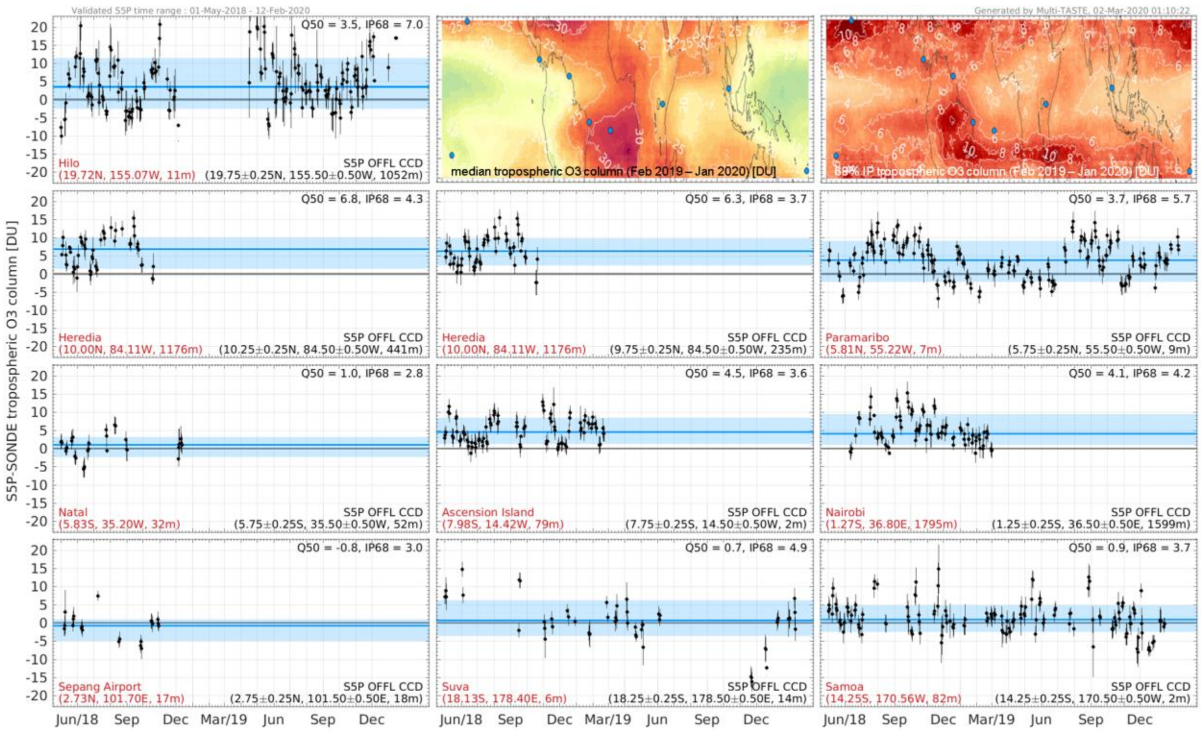


Figure 16: Time series of the absolute difference between spatially and temporally co-located S5P and ozonesonde tropospheric O_3 column data. The blue line and shaded area shows the median value and the range between the 16% and 84% percentiles. Positive values indicate a high bias of S5P w.r.t. the reference. The first row also maps the location of the ozonesonde sites and the characteristics of the tropospheric O_3 field (median and 68% interpercentile over one year of S5P data, see **Figure 14**).

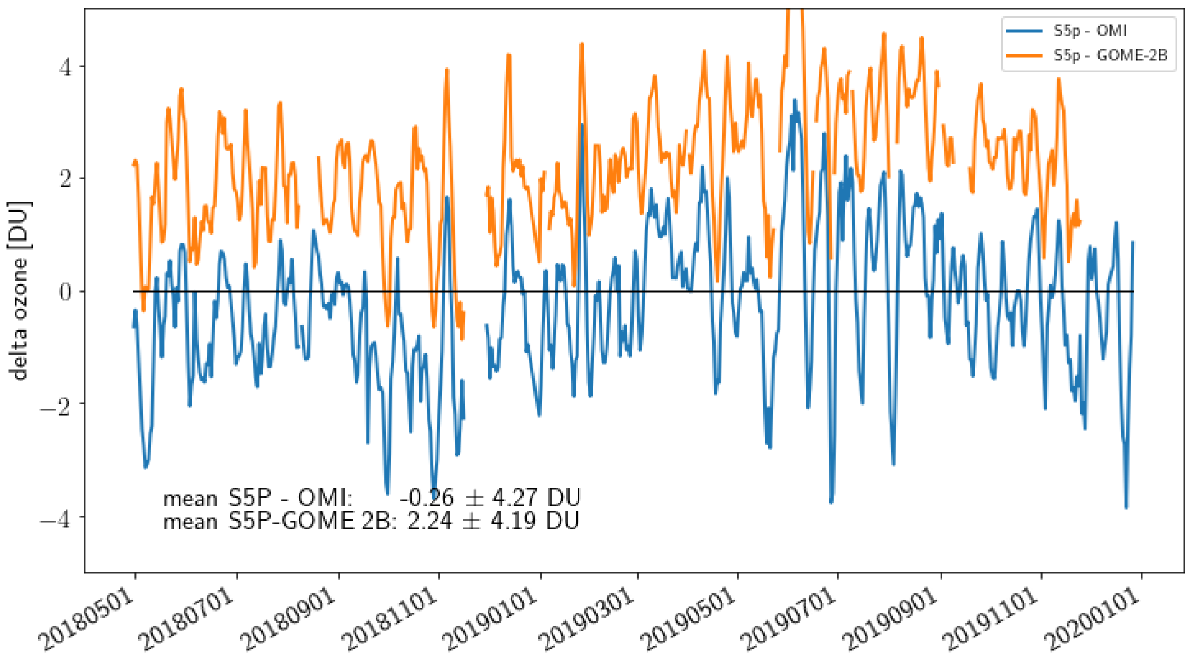


Figure 17: Difference time series of daily tropospheric O_3 column data averaged over the 20°S – 20°N tropical belt. S5P OFFL CCD data are compared to satellite data by OMI and GOME-2B, positive values indicate a high bias of S5P w.r.t. the reference.

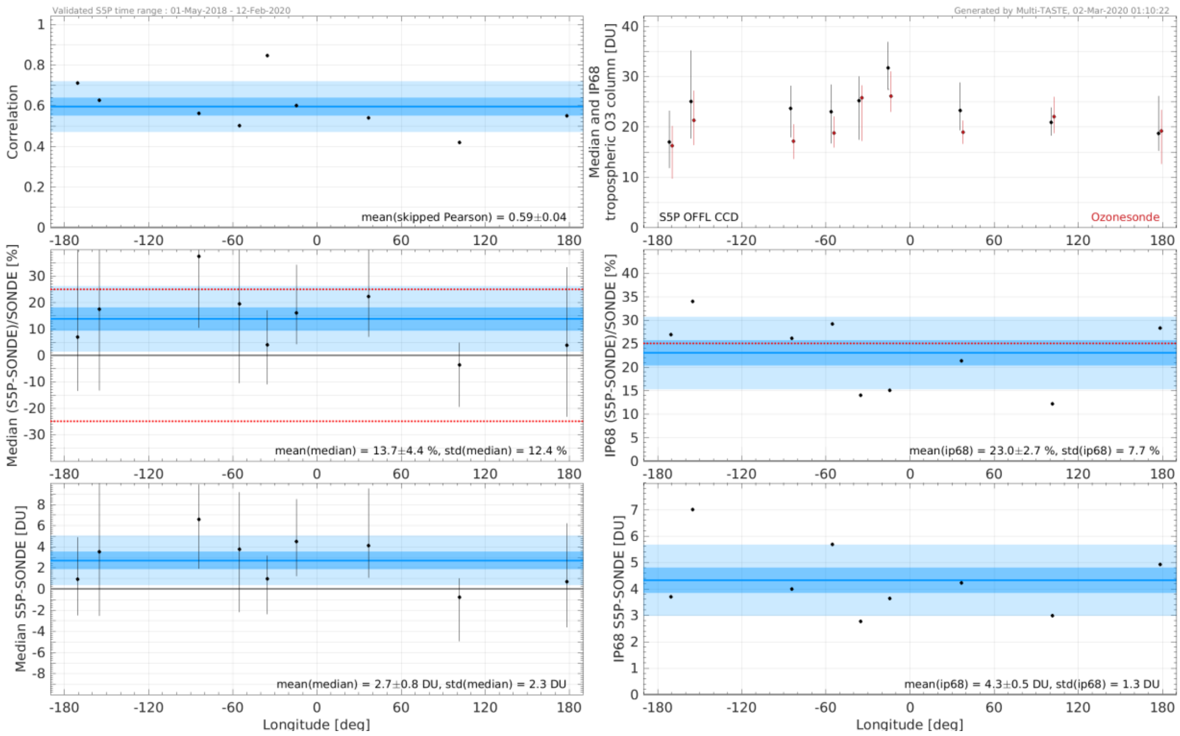


Figure 18: Overview of correlation (top left), median bias (middle & bottom left) and intercomparison spread (middle & bottom right) of S5P tropospheric O_3 column data for each SHADOZ site (black markers). Black vertical bars represent the 68% interpercentile of the comparison time series. The mean, standard error of the mean (1σ) and standard deviation (1σ) of the quality indicator across the network are shown as a horizontal blue line and shaded areas.

6 Validation Results: L2_NO2

6.1 L2_NO2 products and requirements

This Section reports on the validation of the following geophysical variables of the S5P TROPOMI L2_NO2 data products identified in **Table 1**: the NO_2 stratospheric column, the NO_2 tropospheric column, and the NO_2 total column. Validation results are discussed with respect to the product quality targets outlined in **Table 3**.

The NRTI and OFFL processors are producing very similar results. Therefore, mainly the validation of the L2_NO2 OFFL product is reported hereafter. The OFFL product uses ECWMF analysis meteorological data as input for the CTM, while the NRTI uses ECWMF forecast meteorological data. Furthermore, the OFFL product is reprocessed to version 01.02.02. Subsection 6.4 demonstrates evidence that NRTI and OFFL data do not differ significantly and that their respective validations yield similar conclusions.

6.2 Validation approach

6.2.1 Ground-based monitoring networks

Stratospheric NO_2 – ZSL-DOAS UV-Visible Spectrometers

S5P TROPOMI L2_NO2 stratospheric nitrogen dioxide column data are compared routinely to reference measurements acquired by Zenith-Scattered Light Differential Optical Absorption Spectroscopy (ZSL-DOAS) UV-Visible spectrometers (Pommereau and Goutail, 1988; Hendrick et al., 2011). Those instruments perform network operation in the context of the Network for the Detection of Atmospheric Composition Change (NDACC). ZSL-DOAS column data have a bias of less than 10% and a dispersion better than 1%.

To account for effects of the photochemical diurnal cycle of stratospheric NO_2, the ZSL-DOAS measurements, obtained twice daily at twilight at each station, 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. It 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 NO_2 values (from ground pixels at high resolution) are averaged over the footprint of the air mass to which the ground-based zenith-sky measurements are sensitive.

At this stage of the S5P routine operations, most of the ZSL-DOAS validation data have been obtained through the SAOZ near-real-time processing facility operated by the CNRS LATMOS, from 14 stations located between 79°N and 75°S. These are highlighted in red in **Figure 19**. They are now complemented with measurements from 8 other NDACC affiliated ZSL-DOAS instruments yielding co-locations with the S5P data sets.

Tropospheric NO_2 – MAX-DOAS UV-Visible Spectrometers

S5P TROPOMI L2_NO2 tropospheric nitrogen dioxide column data are routinely compared to reference measurements acquired by MAXDOAS (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).

At the present stage of S5P routine operation, four MAX-DOAS stations have contributed data (Athens and Bremen from IUP-B and Cabauw and De Bilt from KNMI) and are included in the operational validation. MAXDOAS tropospheric NO_2 column data have a bias of maximum 20% and a precision better than 30% at this set of stations.

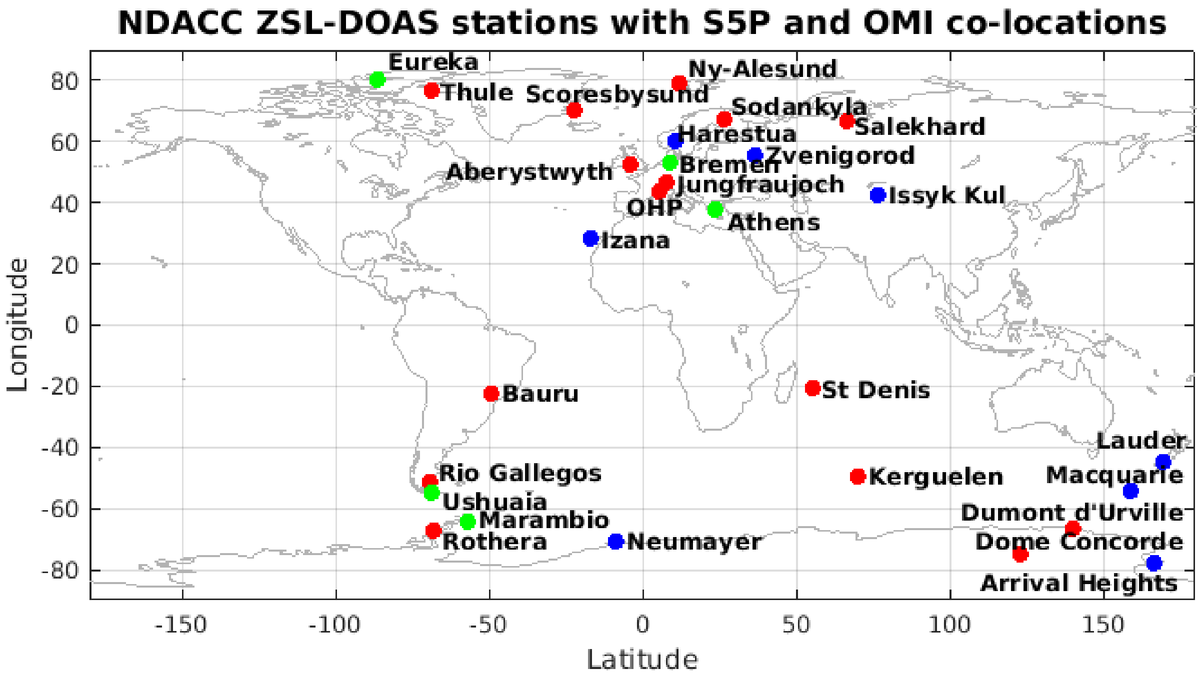


Figure 19: Geographical distribution of the NDACC ZSL-DOAS instruments measuring routinely stratospheric NO_2 and yielding co-locations with the current S5P L2_NO2 data sets. Stations marked with a red dot contribute fast delivery data thanks to the LATMOS_RT facility. Blue and green dots depict other NDACC stations contributing ZSL-DOAS data directly through the NDACC DHF and the AO project NIDFORVAL, respectively.

Total NO_2 – 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 the Pandora system. Those instruments 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%). Comparisons at 14 sites are operational on the VDAF Automated Validation Server. The comparison criteria on the VDAF-AVS are:

- S5P data have a qa_value > 0.5;
- the TROPOMI ground pixel contains the Pandora station;
- any Pandora measurement with a flag not equal to 0 or 10 is excluded, as well as negative Pandora values;
- 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 NO_2 can also be representative of the stratospheric NO_2 column.

6.2.2 Satellites

S5P TROPOMI L2_NO2 nitrogen dioxide column data are also compared to similar 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 those TROPOMI pixels 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. As clouds are less of a problem for S5P stratospheric NO2 retrievals and stratospheric data comparisons, data with `qa_value` above 0.5 are nevertheless used hereafter. 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

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). Results summarized in this section take into account the outcome of the third *Copernicus Sentinel-5 Precursor Validation Team Workshop* held at ESA/ESRIN in Frascati (Rome, Italy) between 11 and 14 Nov 2019, where 10 talks and 8 posters were presented during the NO_2 validation session. These results basically support the findings reported by the routine validation.

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 currently 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.

6.3.3 Stratospheric NO_2 column

6.3.3.1 Bias

The validation results described here are based on comparisons at 19 NDACC ZSL-DOAS stations, sampling the latitude range from 80°N (Eureka) to -75°S (Dome C). The ZSL-DOAS stratospheric NO_2 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.

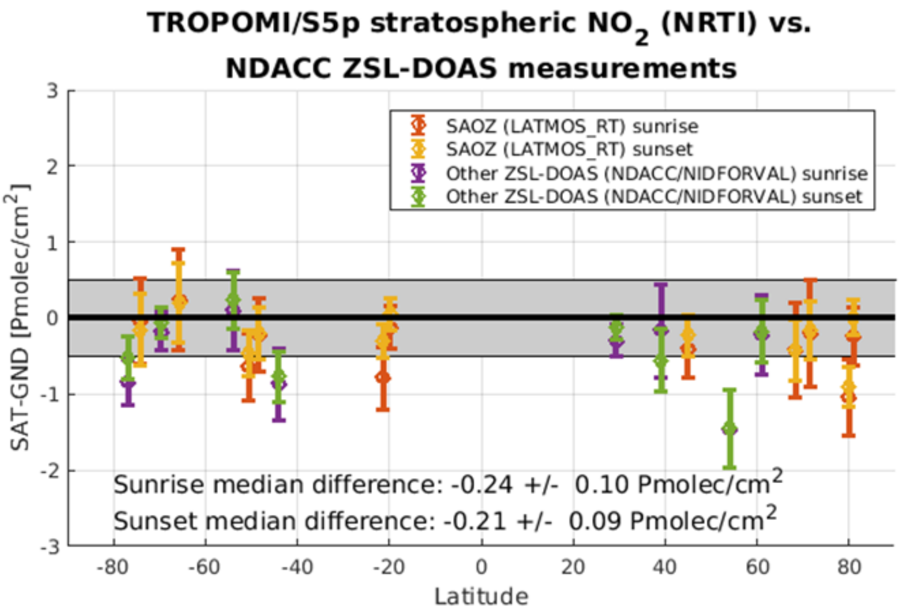


Figure 20: Meridian dependence of the median (the circular markers) and spread (±1σ error bars) of the differences between S5P TROPOMI L2_NO2 (NRTI) stratospheric column data and ZSL-DOAS reference data (of the SAOZ type and other UV-visible instruments), represented at individual 21 NDACC stations from the Antarctic to the Arctic. The time frame is the phase E2 (May 2018) until February 2020. The values in the legend correspond to the network-wide mean bias and its formal uncertainty. This graph includes results for processor versions greater or equal to 1.1.0.

S5P L2_NO2 NRTI stratospheric column data are generally lower than the ground-based values by approximately -0.2 Pmolec/cm^2, with a station-to-station scatter of this bias of similar magnitude (**Figure 20**). This is within the mission requirement of a maximum bias of 10% (equivalent to 0.2-0.4 Pmolec/cm^2, depending on latitude and season). The L2_NO2 OFFL median bias is -5.8% (-0.18 Pmolec/cm^2) with a station-to-station scatter of 10%, which is also within the bias requirement of 10%.

Three of the Pandora instruments are at elevated sites: Altzomoni (3985 m),Izaña (2360 m), Mauna Loa (4169 m), where the measured signal corresponds more to the S5P L2_NO2 stratospheric column rather than the total column (see **Figure 21**). Comparison data was obtained from the VDAF-AVS on 2019-11-25 and the negative Pandora outliers removed.

Figure 21 presents boxplots of the difference and of the relative difference between S5P L2_NO2 NO_2 stratospheric column data and Pandora total column data. The median relative difference is within the 10% bias requirement. It should be noted that, although the Pandora instrument at these sites usually are above low-lying tropospheric pollution, its measurement remains sensitive to NO_2 in the free troposphere, which could introduce a small positive bias in the comparison.

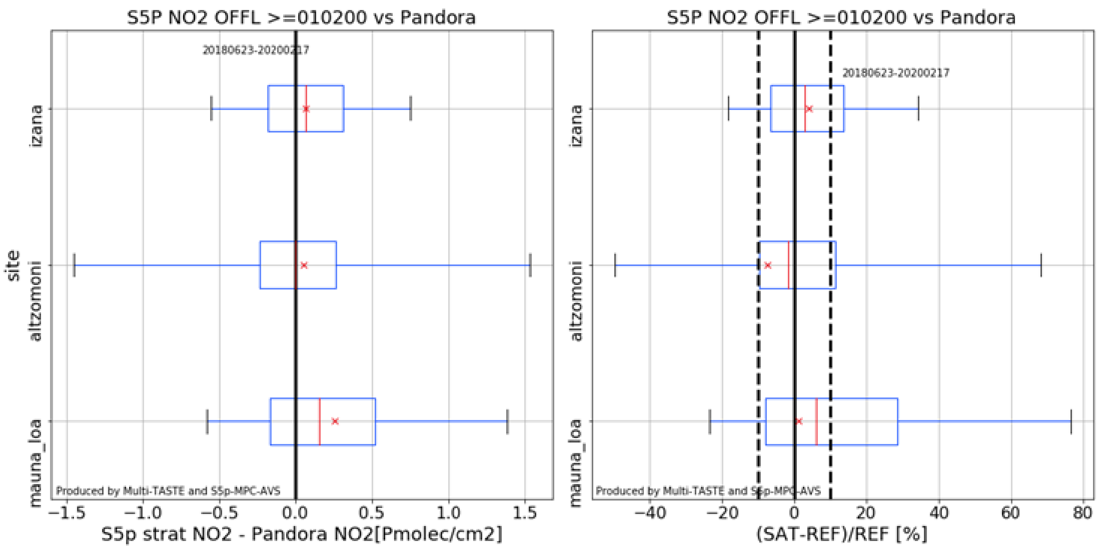


Figure 21: Boxplots of S5P RPRO+OFFL stratospheric NO_2 column vs. PGN Pandora total NO_2 column at three mountain sites: difference (left) and relative difference (right). Data was obtained from the VDAF Automated Validation Server on 27/02/2020. Apart from the filters in the AVS (S5P: qa_value > 0.5, Pandora: only quality flag values of 0 and 10 are kept), negative Pandora values were also removed. Boxplot conventions: box bounds are at first and third quartile. Red line is the median. Whiskers are at 5% and 95% percentile. Red cross is the mean difference.

6.3.3.2 Dispersion

The ±1σ dispersion of the difference between S5P 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 20**). 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 max. 0.5 Pmolec/cm². The mean scatter of the OFFL L2_NO2 is 0.33±0.18 Pmolec/cm².

Considering the comparison of S5P stratospheric NO_2 VCD vs. PGN Pandora total NO_2 at the three elevated sites (see **Figure 21**) the standard deviation ranges from 0.4 Pmolec/cm^2 (Izaña) to 0.9 Pmolec/cm^2 (Altzomoni), while ½ IP68 ranges from 0.4 Pmolec/cm^2 (Izaña) to 0.5 Pmolec/cm^2 (Altzomoni and Mauna Loa). Given that the Pandora measurements display some scatter, we infer that the dispersion requirement of 0.5 Pmolec/cm² is satisfied.

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.cm^{-2} over the range of those influence quantities (**Figure 22**).

6.3.3.4 Seasonal cycle and short term variability

S5P and ground-based ZSL-DOAS instruments capture similarly the short-term variability (at daily and monthly scales) of the NO_2 stratospheric column, as illustrated at the NDACC station of Kerguelen Island in **Figure 23**. The ground-based SAOZ data acquired at twilight were adjusted to account for the photochemical diurnal variation between twilight and the S5P overpass time.

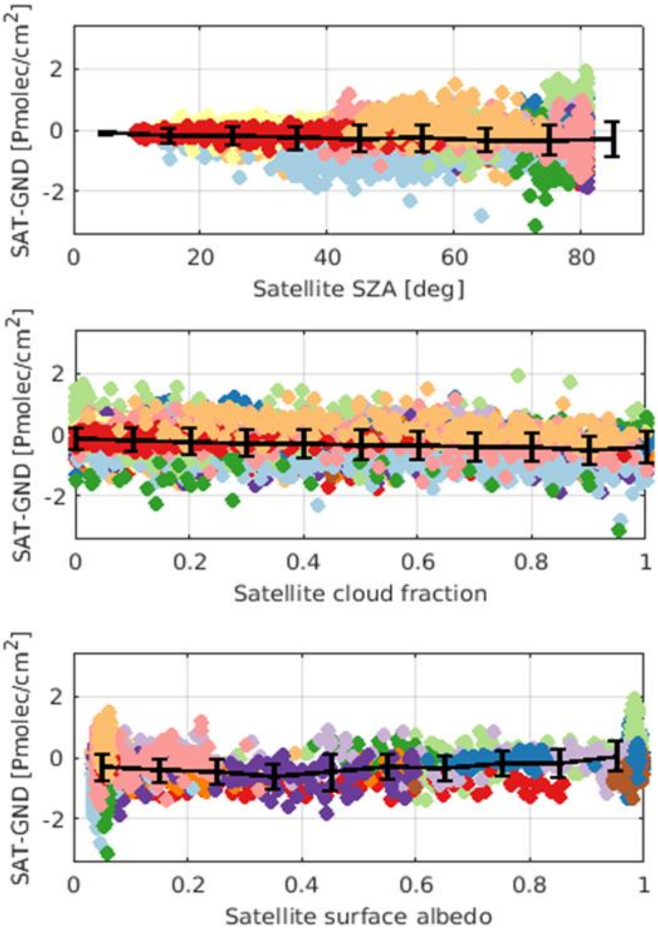


Figure 22: Difference between S5P L2_NO2 NRTI and ground-based SAOZ stratospheric NO_2 column data as a function of the satellite solar zenith angle (SZA), satellite cloud fraction and satellite surface albedo. Mean and standard deviation calculated over bin widths of 10 degrees in SZA, 0.1 in CF, and 0.1 in surface albedo.

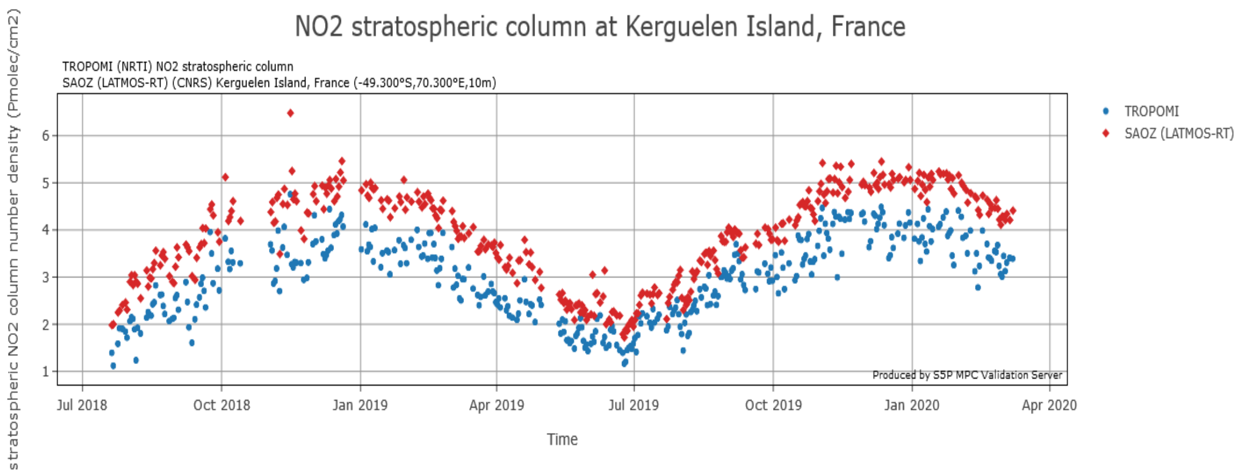


Figure 23: Time series of S5P NRTI L2_NO2 stratospheric NO_2 column data co-located with ground-based SAOZ twilight measurements performed by LATMOS at the NDACC southern mid-latitude station of Kerguelen Island. Photochemical correction is deactivated to offset the two time series and better see the day-to-day variability.

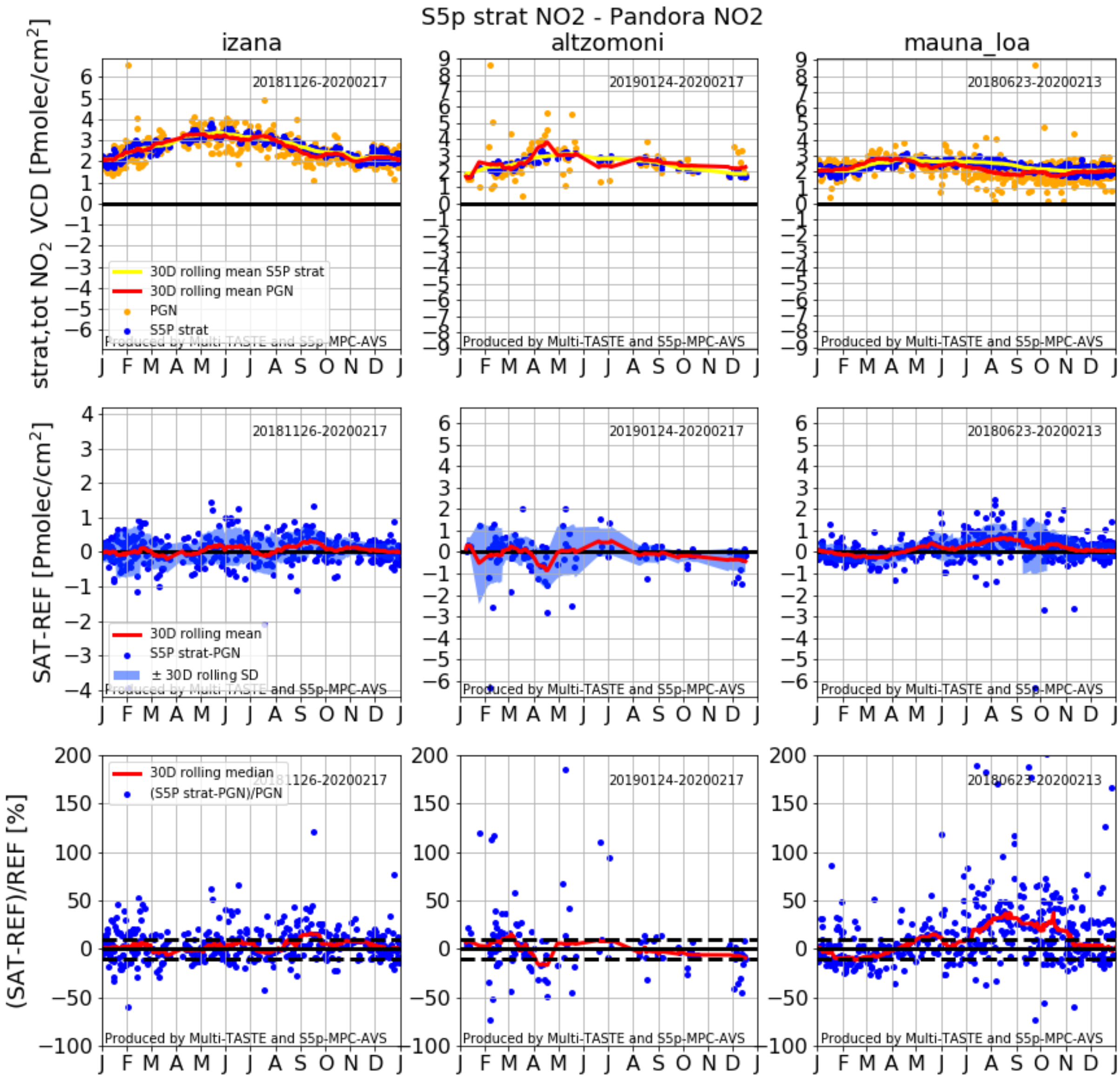


Figure 24: Seasonal cycle plots of the S5P stratospheric NO_2 column and Pandora total column column at 3 mountain sites. Upper: S5P and PGN Pandora together; middle: their difference; lower: their percent relative difference. Plain lines represent the 30-day rolling mean or (for relative difference) the 30-day rolling median (cyclic boundary conditions).

S5P stratospheric NO_2 and Pandora total NO_2 plotted as a function of the calendar month at 3 mountain sites follow the same seasonal cycle (**Figure 24**). 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 July-November at Mauna Loa, which this seems to be due to low-lying Pandora values

6.3.3.5 Geographical patterns

None to report.

6.3.3.6 Switch to smaller ground pixel resolution

The effect of the change in TROPOMI ground pixel size on 6 August 2019 on the S5P stratospheric NO_2 column data was investigated by comparing the S5P and ground-based time series at the NDACC ZSL-DOAS stations (**Figure 25)** and 3 PGN Pandora mountain sites (**Figure 26**). The difference between S5P and the ground-based data does not show any impact of the pixel size change.

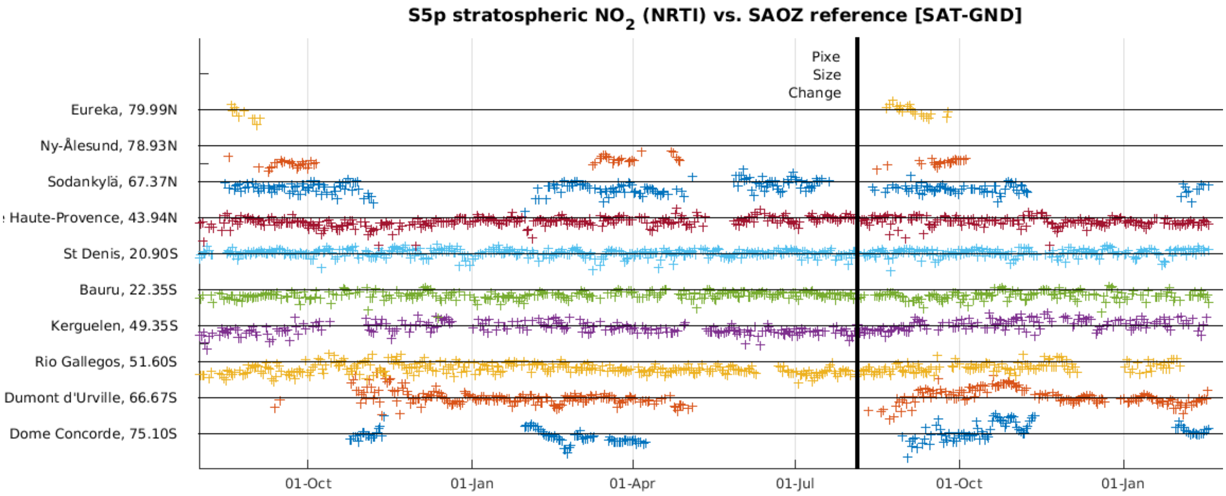


Figure 25: Time series – from August 2018 until February 2020 – of the difference between S5P NRTI L2_NO2 and NDACC SAOZ NO_2 stratospheric column data. Solid horizontal lines are spaced by 1 Pmolec.cm^{-2}. The black line indicates the TROPOMI switch to finer horizontal resolution.

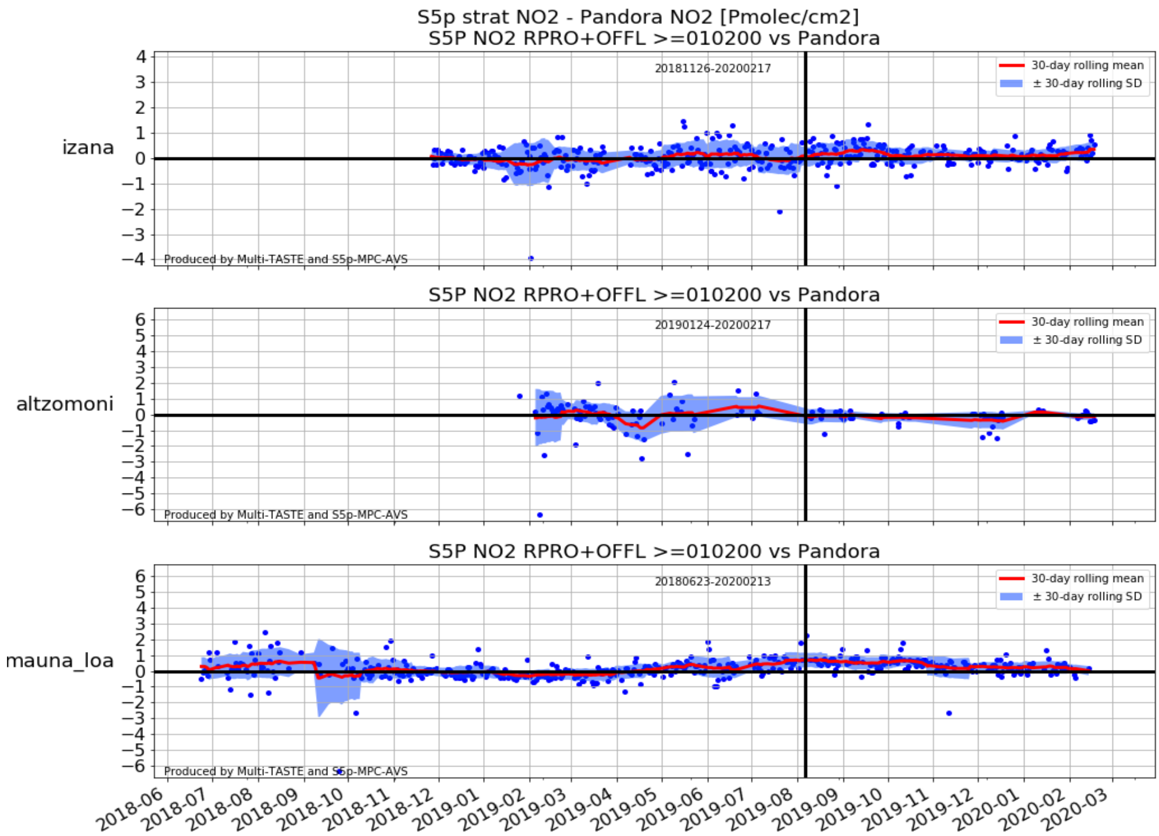


Figure 26: Time series from June 2018 until February 2020 of the difference between S5P RPRO+OFFL L2_NO2 NO_2 stratospheric column data and PGN Pandora total NO_2 data at 3 mountain sites.

6.3.3.7 *Other features*

None to report.

6.3.4 Tropospheric NO_2 column

6.3.4.1 *Bias*

S5P L2_NO2 RPRO+OFFL NO_2 tropospheric column data have been compared to the ground-based MAX-DOAS column data at 4 stations in Europe using the VDAF Automated Validation Server (inspection at 2020-02-27), yielding between 272 to 376 colocations per site, and 1265 measurement pairs. The mean difference at each site varies between -1.4 Pmolec/cm² and -2.7 Pmolec/cm² and the median difference between -0.2 Pmolec/cm² and -2.2 Pmolec/cm² (see **Figure 27**). The median relative difference varies between -10% (Athens) and -30% (Cabauw, Bremen). The median difference calculated over all comparison pairs is -23%. Comparisons with ground-based data at 10 Pandora and 17 MAXDOAS stations collected through the AO project NIDFORVal show similar results. Also, a negative bias within the -50% requirement is found. This bias is reduced by several percent when vertical averaging kernels are applied to reduce smoothing difference errors.

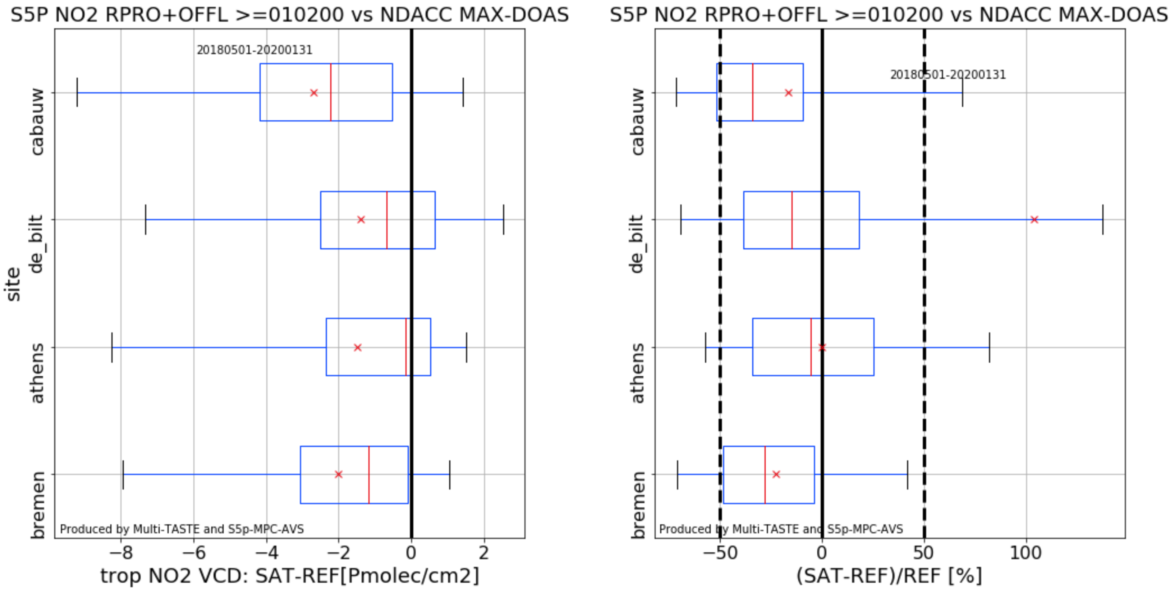


Figure 27: Comparison of S5P RPRO+OFFL vs. MAX-DOAS tropospheric NO_2 column data at four European stations. Difference (left) and relative difference (right). Data was obtained from the VDAF Automated Validation Server on 27/02/2020. Boxplot conventions: box bounds are at first and third quartile. Red line represents the median difference while red crosses represent the mean difference. Whiskers are at 5 and 95 percentiles.

6.3.4.2 Dispersion

The standard deviation of S5P tropospheric column data w.r.t. MAX-DOAS varies between 3 and 4 Pmolec.cm^{-2}. This exceeds by far the mission precision target of 0.7 Pmolec.cm^{-2}, however, 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 is approximately 2 Pmolec/cm^2. There is a reasonably good correlation between S5P tropospheric column and MAX-DOAS data, with the Pearson R varying between 0.63 (Bremen) and 0.78 (Athens).

6.3.4.3 Dependence on influence quantities

None to report.

6.3.4.4 Seasonal and shorter term variability

Figure 28 presents the seasonal cycle of the difference between S5P RPRO+OFFL and NDACC MAX-DOAS tropospheric NO_2 (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.

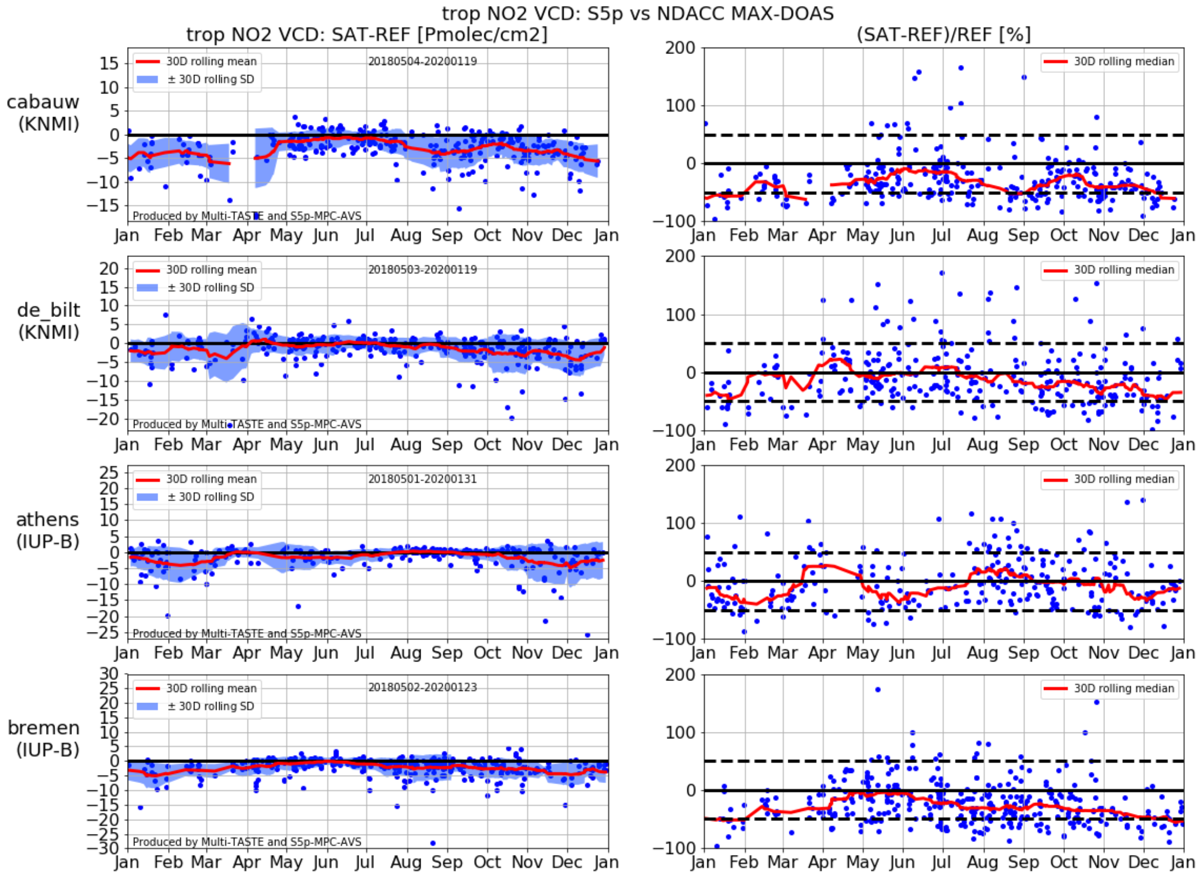


Figure 28: Seasonal cycle of the difference between S5P RPRO+OFFL and MAX-DOAS NO_2 tropospheric column data at four European stations. Difference (left) and relative difference (right). Data was obtained from the VDAF Automated Validation Server on 27/02/2020.

S5P reports lower values than MAX-DOAS in late fall and in wintertime, when tropospheric NO_2 reaches its largest abundance. Over the entire year the 30-day rolling median relative difference is within the mission requirements for the bias.

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 29**).

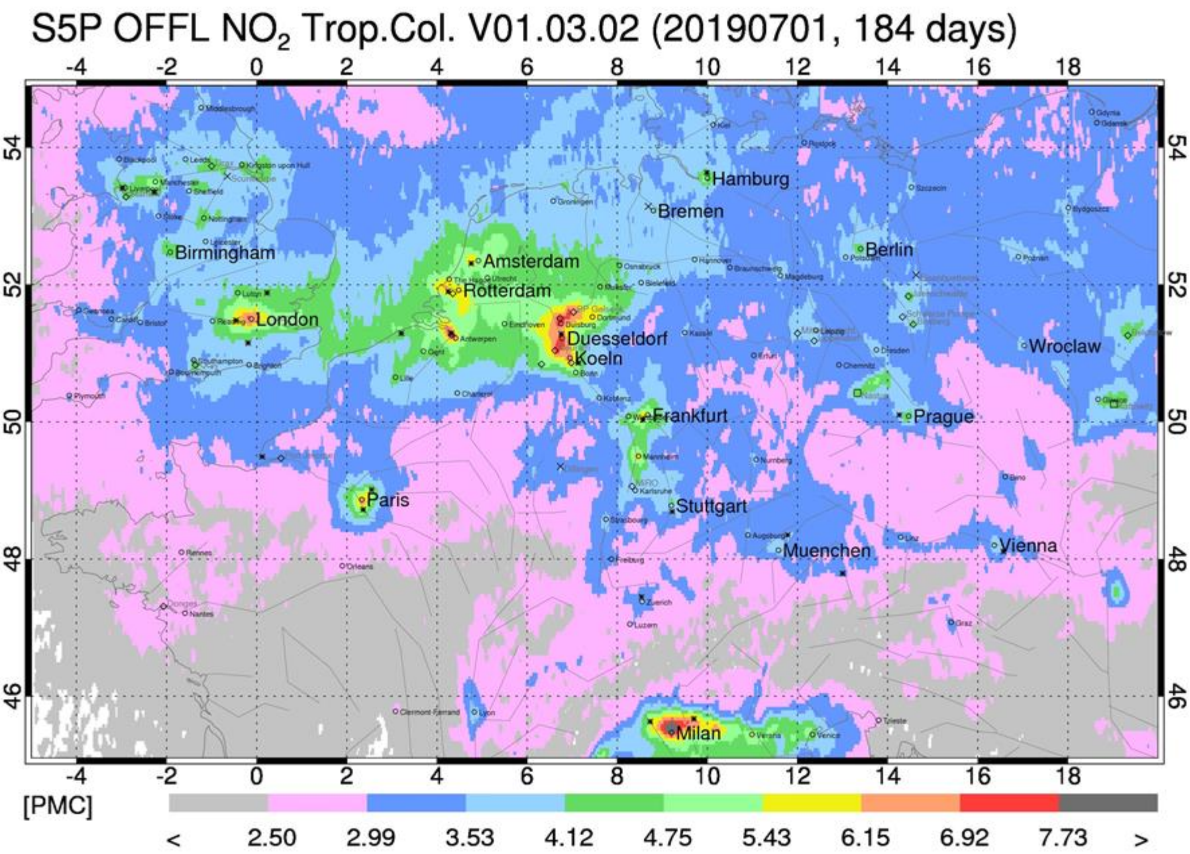


Figure 29: S5P tropospheric NO_2 over central Europe. Processor versions OFFL 01.03.00-02 was used to bin the data on a 0.03°x0.06° grid for a 6-month period starting July 2019. The qa_value flags larger than 0.5 and a CF<0.6 were chosen to reduce the amount of data and exclude cloudy scenes. 1 PMC means 1 Pmolec/cm^2.

6.3.4.6 Switch to smaller ground pixel resolution

None 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 (0.60-0.82, see before) but low slopes. In S5P=a*MXD+b, a varies between 0.3 (Bremen) 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*MXD+b, where a varies now between 0.71 (Athens) and 0.90 (De Bilt).

6.3.5 Total NO_2 column

6.3.5.1 Bias

Based on measurements from 19 Pandora stations (**Figure 30**) between 64.9°N and -35.3°S, the overall median bias (5628 measurement pairs, 2020-02) is -8.0% with a high station-to-station scatter of 26%. These results are now well within the accuracy requirements of 30%, which is the average of the tropospheric and stratospheric bias maxima.

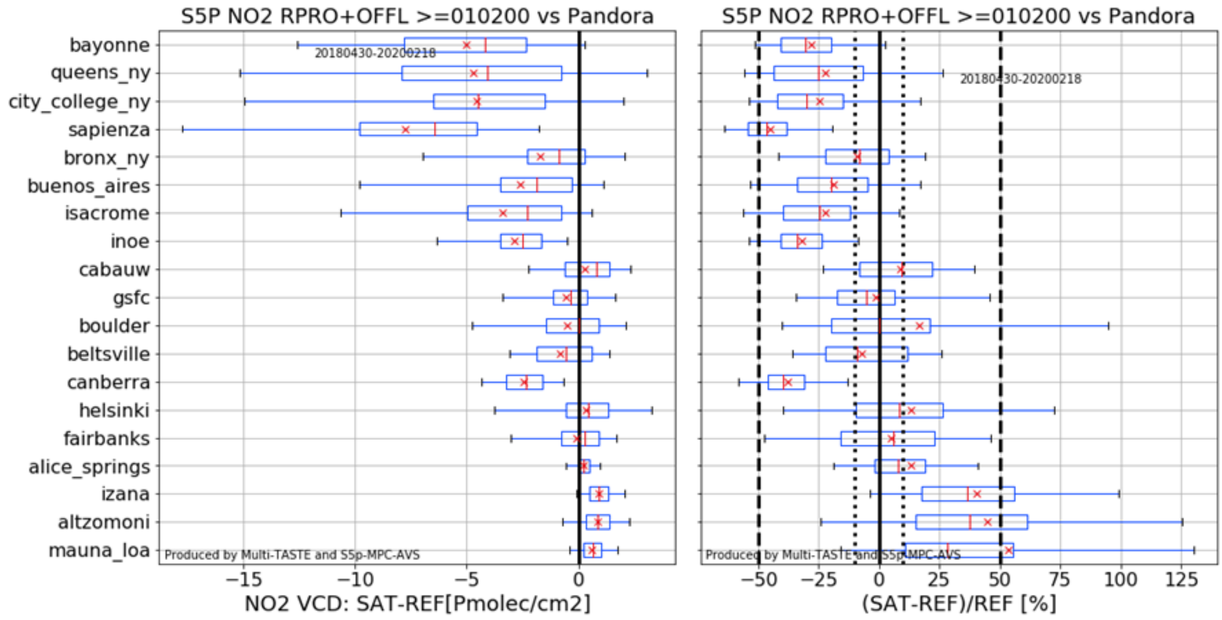


Figure 30: Boxplots of S5P RPRO+OFFL total NO_2 column vs. PGN Pandora total NO_2 column. Difference (left) and relative difference (right). Sites are ordered according to mean Pandora total NO_2 column value (highest at top); note the three mountain sites are at the bottom. Data was obtained from the VDAF Automated Validation Server at 2020/02/28. It covers the time frame from May 2018 to February 2020. Note that regarding Pandora data, only data with flags 0 and 10 are kept. Furthermore, negative Pandora values were removed. Regarding S5P data, the filter qa_value>0.5 was applied if tropospheric NO_2 /stratospheric NO_2 ratio < 1 and qa_value>0.75 was applied otherwise. Boxplot conventions: box bounds are at first and third quartile. Red line is median. Whiskers are at 5 and 95 percentiles. Red cross is mean.

We highlight here 3 different comparison cases. At Alice Springs (Australia), where NO_2 column values are small (mostly between 2-4 Pmolec/cm^2), a small positive bias of 0.2 Pmolec/cm^2 is seen (+8% median relative difference). At New York Bronx (United States), there is a wider distribution of NO_2 values (2-30 Pmolec/cm^2). The bias is negative (mean difference = -3 Pmolec/cm^2, median relative difference = -15%). Finally, at Sapienza (Rome, Italy), Pandonia column values can reach almost 40 Pmolec/cm^2. S5P displays a bias (mean difference) of -8 Pmolec/cm^2 or -47% median relative difference here. Locally enhanced NO_2 probably contributes to this.

6.3.5.2 Dispersion

The dispersion of the difference between S5P and PGN Pandora measurements depends strongly on the station. Small standard deviations are observed at Alice Springs, Izaña and Mauna Loa (0.6-0.7 Pmolec/cm^2 i.e., comparable to the mission precision requirement), and higher values elsewhere (e.g., 6 Pmolec/cm^2 at New York Bronx and at Sapienza Rome, 8 Pmolec/cm^2 at New York City College). The site averaged standard deviation is 3.4±2.4 Pmolec/cm^2. The Pearson-R varies from relatively low (e.g., 0.4-0.5 at Fairbanks and Helsinki) to high (0.86 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 *Switch to smaller ground pixel resolution*

The effect of the change in TROPOMI ground pixel size on 6 August 2019 on the S5P total NO_2 data was investigated by having a close look at the S5P and PGN Pandora time series at every individual station separately. **Figure 31** shows that the bias and scatter of the difference are not affected by the pixel size change.

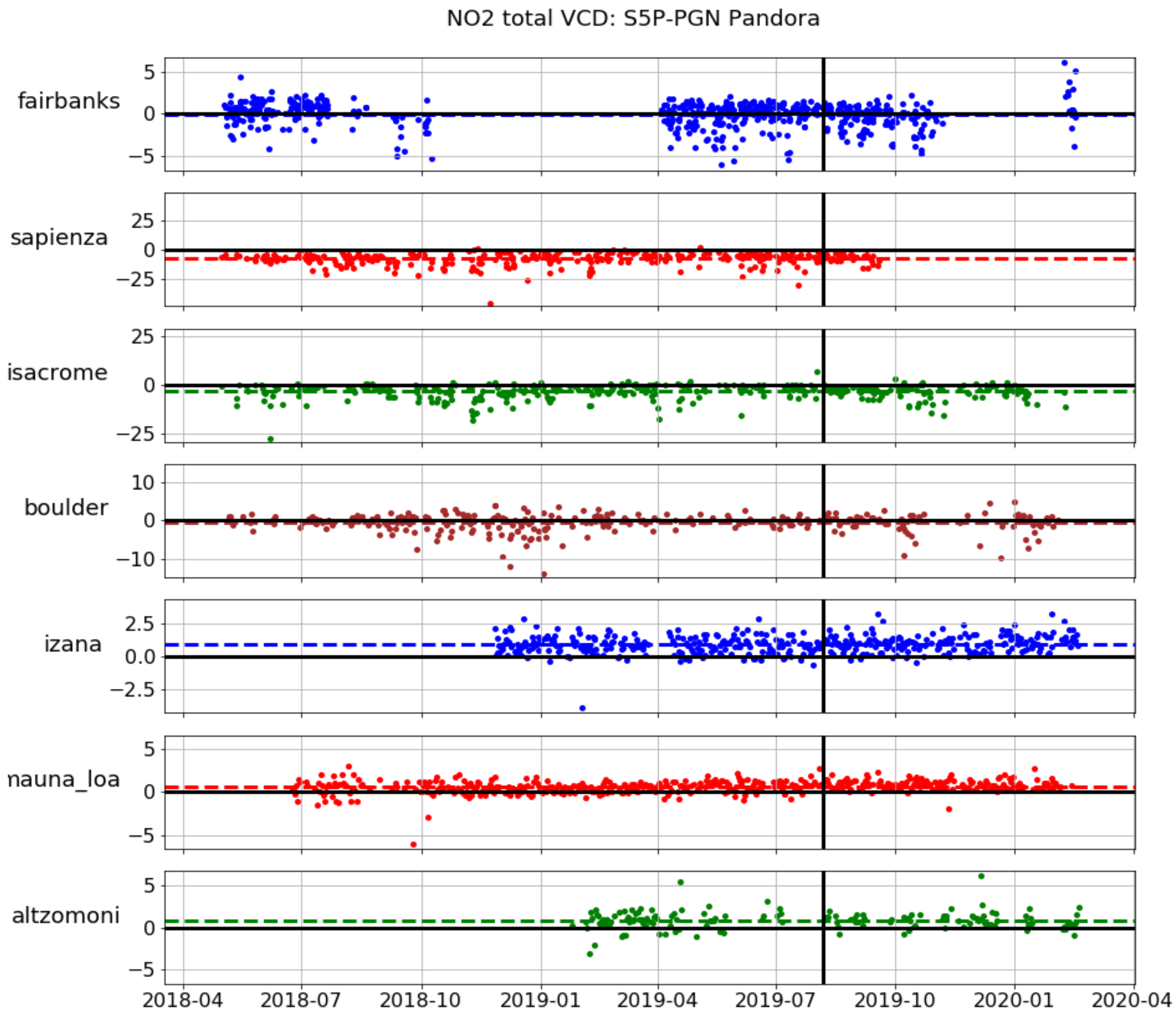


Figure 31: Time series of the difference between S5P RPRO+OFFL and ground-based PGN Pandora NO_2 total column data, from June 2018 until February 2020. The black line indicates the S5P switch to finer horizontal resolution.

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 NO_2 Column

The similarity of the two products can be investigated by comparing the processing of a randomly chosen orbit. **Figure 32** shows this approach for orbit 7407 on March 19, 2019. It reveals differences mostly below the mission requirement on precision (0.5 Pmolec/cm^2). The RMSD is 0.16 Pmolec/cm^2, with values up to 0.5 Pmolec/cm^2. Some features are due to a different stratosphere/troposphere division (e.g. north-east of Iceland, positive difference in stratosphere, negative in troposphere).

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.

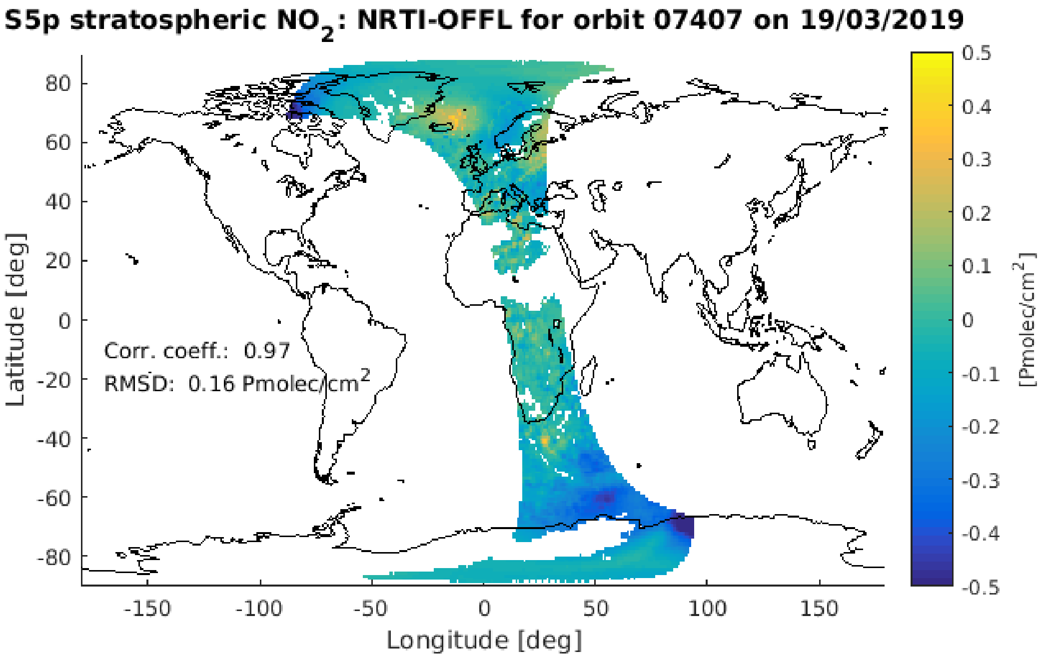


Figure 32: Difference between S5P NRTI and OFFL stratospheric NO_2 column data for a single orbit (gridded to 1°x1° resolution).

The resulting OFFL pole-to-pole graph is shown in **Figure 33**, illustrating that in the end, OFFL performs very similarly to NRTI with a bias of -0.2 Pmolec/cm^2.

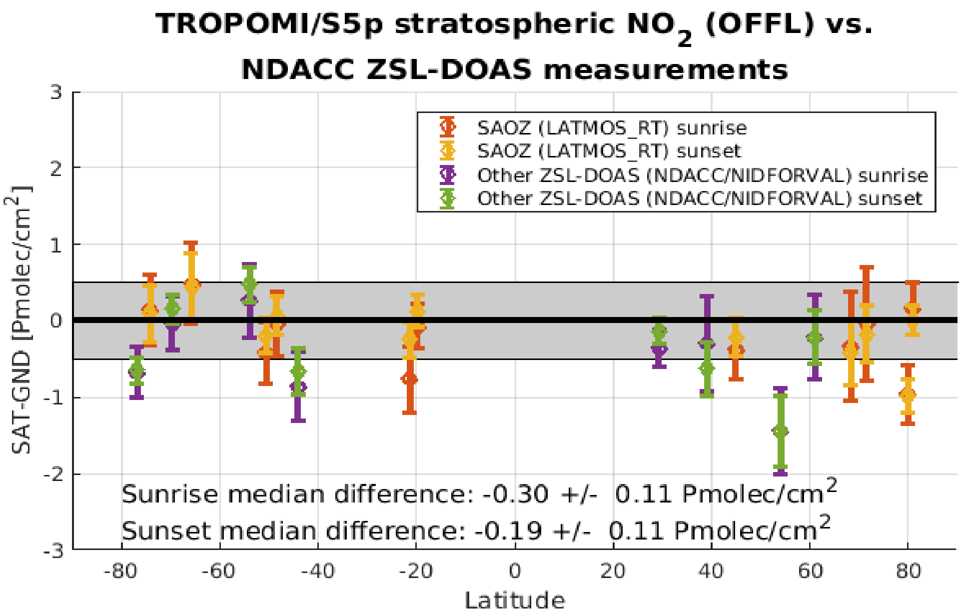


Figure 33: Meridian dependence of the mean (the circular markers) and dispersion (±1σ error bars) of the differences between S5p TROPOMI L2_NO2 (OFFL) stratospheric column data and ZSL-DOAS reference data, represented at individual stations from the Antarctic to the Arctic. The values in the legend correspond to the median and its formal uncertainty for all mean (per station) differences.

6.4.2 Tropospheric NO2 Column

To demonstrate the closeness of L2_NO2 NRTI and OFFL products at the MAX-DOAS sites Athens, Bremen, De Bilt and Cabauw, 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. The subset of pixels common to both NRTI and OFFL were determined and compared. Differences between NRTI, OFFL and MAX-DOAS were determined. Statistical results for Athens and Bremen are summarized **Table 5**: similar conclusions on the closeness of NRTI and OFFL are obtained for the sites De Bilt and Cabauw.

Table 5 – Statistics on the comparison of the common subset of L2_HCHO NRTI, L2_HCHO RPRO+OFFL and co-located MAX-DOAS, for the sites Bremen and Athens. (*: unit of Pmolec/cm^2). Numbers updated using data from the validation server on 2020-02-27.

	Bremen 244 co-locations			Athens 242 co-locations		
	NRTI-OFFL	NRTI-MXD	OFFL-MXD	NRTI-OFFL	NRTI-MXD	OFFL-MXD
Mean(diff)*	0.11	-2.24±0.23	-2.35±0.23	0.11	-1.43±0.24	-1.54±0.24
Median(diff)*	0.07	-1.39	-1.39	0.10	-0.05	-0.15
Std(diff)*	0.4	3.6	3.6	0.2	3.8	3.8
½ IP68(diff)*	0.3	2.6	2.6	0.13	2.4	2.4
Pearson R	0.98	0.54	0.55	0.998	0.81	0.82
Slope	1.00	0.27	0.27	1.03	0.49	0.48

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 not statistically significant. Also the difference dispersion between NRTI and OFFL is small compared to the difference dispersion between either NRTI or OFFL on 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.

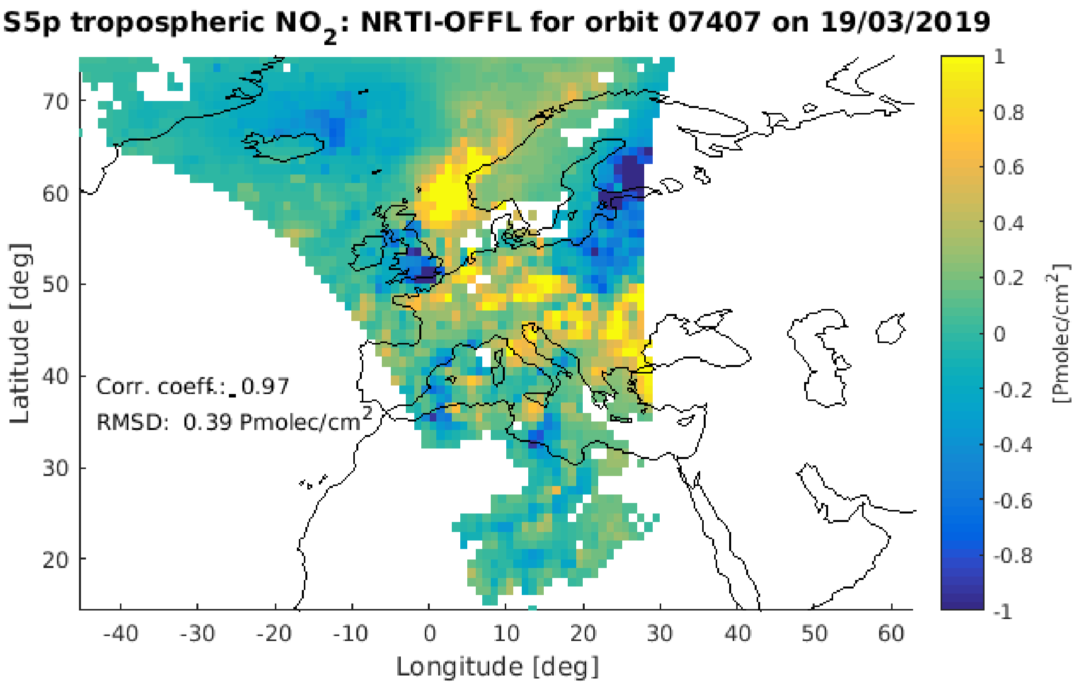


Figure 34: Difference between S5P NRTI and OFFL tropospheric NO_2 column data for a single orbit (gridded to 1°x1° resolution, zoom over Europe).

The tropospheric RMSD is 0.39 Pmolec/cm^2, with values up to 2 Pmolec/cm^2 (**Figure 35**, zoom over Europe). Note strong differences over UK pollution hot spots in this particular orbit (London, Liverpool, Manchester).

6.4.3 Total NO_2 Column

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 35**). The relative difference is in the range of -1%, where NRTI values are slightly higher. A comparison on the single orbit previously analyzed for the stratospheric and tropospheric columns reveals a combination of the features already seen in these subcolumns (**Figure 36**).

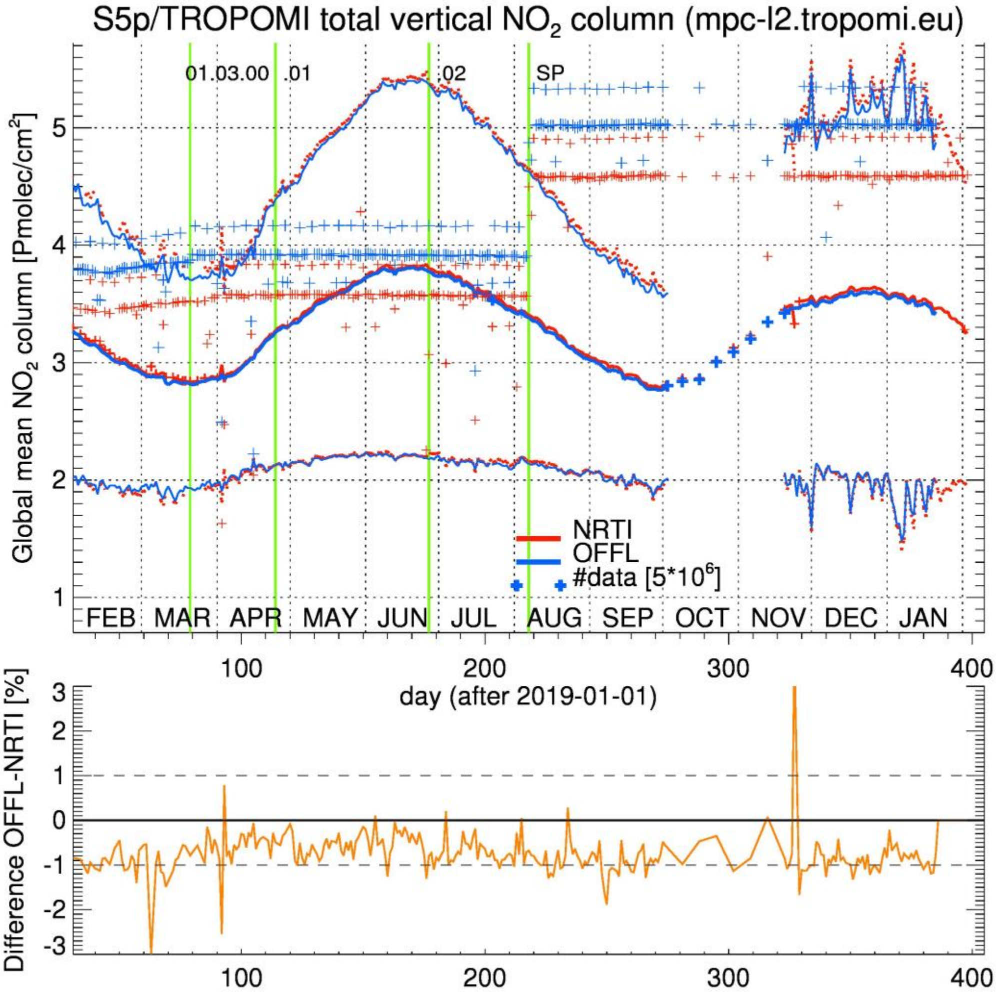


Figure 35: Time series (February 2019 – 2020) of the global mean of NRTI (red) and OFFL (blue) NO2 total column data [Pmolec/cm²]. Data is taken from the TROPOMI QC portal. The ±1σ standard deviations are shown as solid and dotted lines. The number of data points for the NRTI/OFFL data are shown by crosses. The values were divided by a factor of 5*106. Superimposed are the points of processor changes and small pixel switch as green vertical lines. The lower plot shows the difference [%] between OFFL vs. NRTI daily means. Data is taken from the TROPOMI QC portal.

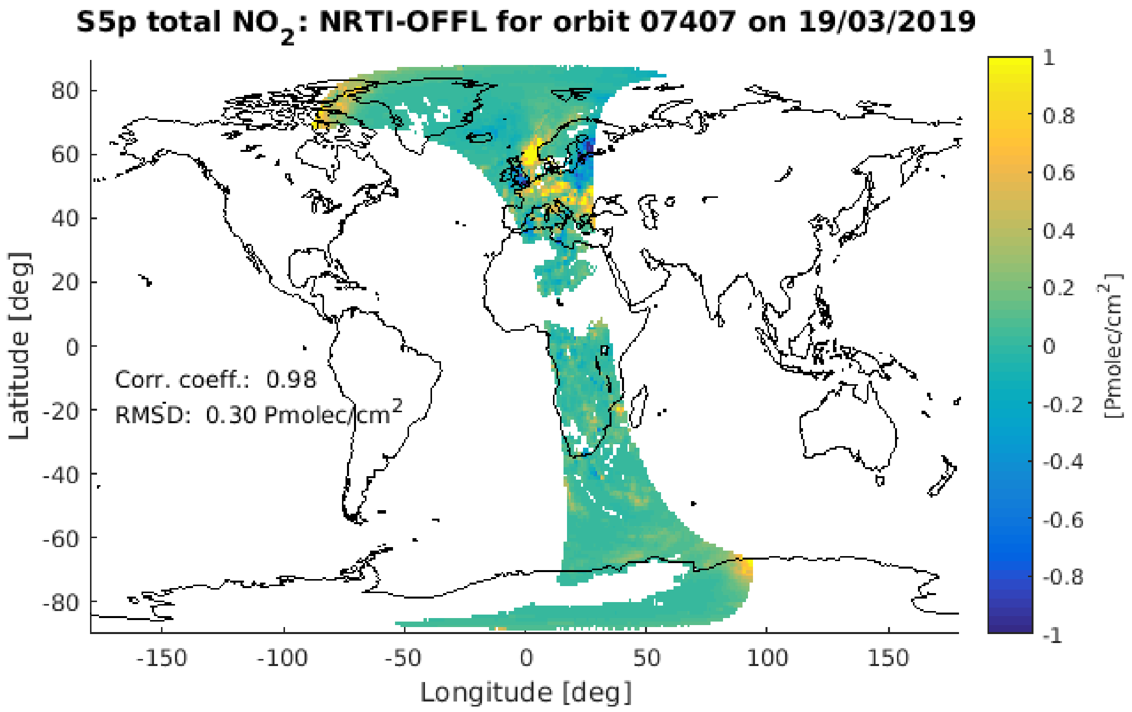


Figure 36: Difference between S5P NRTI and OFFL total NO_2 column data for a single orbit (gridded to 1°x1° resolution).

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, only validation of the L2_HCHO OFFL product is reported hereafter. Subsection 7.4 demonstrates evidence that NRTI and OFFL data do not differ significantly and that their respective validations yield similar conclusions

Notes:

- The operational (E2) phase for the S5P TROPOMI mission starts with orbit #02818.
- The L2_HCHO NRTI and OFFL product have been released in April 2019 with version 01.01.07.
- The L2_HCHO RPRO product 01.01.05 has been released in May 2019, covering the period from 26 June to November 2018.

7.2 Validation approach

7.2.1 Ground-based networks

S5P L2_HCHO data are validated routinely 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 which offer 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 over 25 NDACC FTIR stations. FTIR measurements have a median systematic uncertainty of 13% and a median random uncertainty of 3×10^{14} molec/cm^2 (Vigouroux et al., 2018).

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. At the present stage of S5P routine operation, two MAX-DOAS stations contribute data routinely to the VDAF Automated Validation Server. Seven others provide data through the NIDFORVAL project. MAX-DOAS HCHO column data have a bias of maximum 20% and a precision better than 30%.

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 to EOS-Aura Ozone Monitoring Instrument (OMI). Two versions of the OMI L2 HCHO product are considered (1) the NASA L2 product (10.5067/Aura/OMI/DATA2015), (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; the results can be directly compared because the algorithms have been made as consistent as possible.

7.2.3 Field campaigns and modelling support

Nothing in this report.

7.3 Validation of L2_ HCHO

7.3.1 Recommendations for data usage followed

In order to avoid misinterpretation of the data quality, as recommended, only those TROPOMI pixels associated with a `qa_value` above 0.5 (no error flag, cloud radiance fraction at 340 nm<0.5, SZA<=70°, surface albedo<=0.2, no snow/ice warning, air mass factor>0.1) have been used. 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. Up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal at http://mpc-vdaf.tropomi.eu.

The Copernicus Sentinel-5 Precursor Validation Team Workshop was held at ESA/ESRIN in Frascati (Rome, Italy) between 11 and 14 Nov 2019. 4 talks and 1 poster were presented during the HCHO validation session. These results basically support the findings reported here.

7.3.3 Bias

The following results, using ground-based FTIR and MAX-DOAS, and OMI satellite data sets, show that the S5P L2_HCHO bias usually is within 40%, and always within the upper limit of 80% of the mission requirements.

7.3.3.1 Fourier Transform Infrared Spectrometers

With respect to correlative data at 25 NDACC FTIR stations, S5P L2_HCHO shows a negative bias for high emission sites (-31% for HCHO >8×10^{15} molec/cm^2) and a positive bias for clean sites (+26% for HCHO <2.5×10^{15} molec/cm^2). This is illustrated in **Figure 37**. More details are described in Vigouroux et al. (2020).

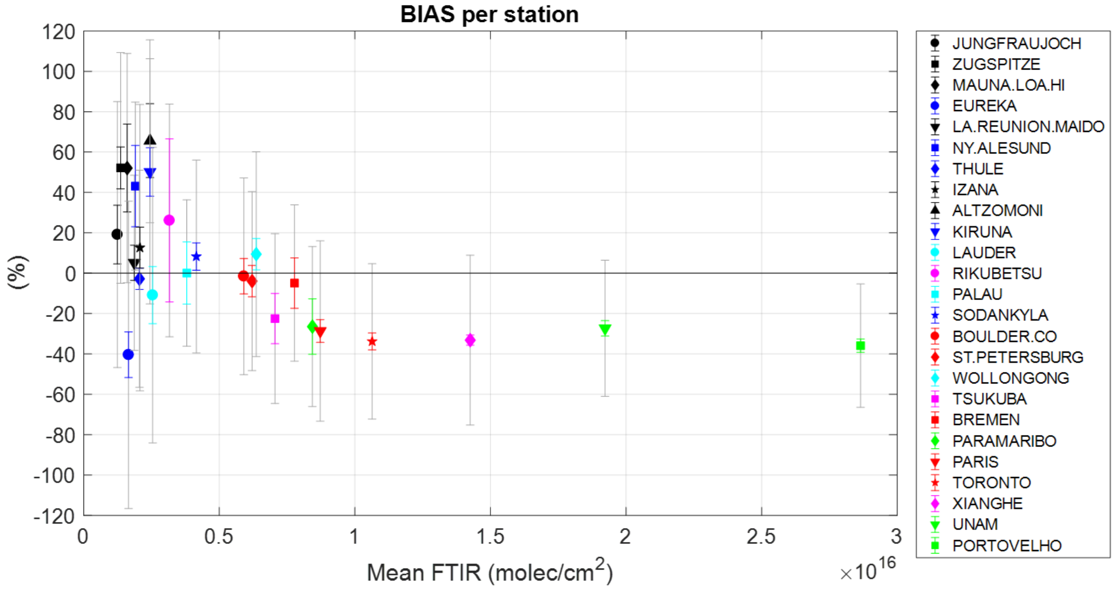


Figure 37: Percent difference between S5P L2_HCHO and FTIR HCHO column data at each station as a function of the mean FTIR total column value (molec/cm2). The gray bars are the systematic uncertainty on the difference, and the colored error bars are the 2-σ error on the bias.

This results point to the presence of both a constant positive bias of TROPMI (1.10±0.05 x10^{15} molec/cm^2), and a proportional one of 0.64±0.03, as obtained using the scatter plot and the robust Theil-Sen estimator to derive the slope and intercept of TROPOMI vs FTIR (see Fig.4 in Vigouroux et al., 2020).

7.3.3.2 MAX-DOAS UV-visible Spectrometers

For direct comparisons of the column data as done in the Automated Validation Server at the time being, the bias of S5P L2_HCHO with respect to MAX-DOAS total HCHO data is -44% (median relative difference) both at Cabauw and De Bilt. This result is within the mission requirement of 80% for the bias.

Further investigation with MAX-DOAS data at 7 other stations provided by the NIDFORVAL AO project, shows that, after mutual application of the averaging kernels of the S5P and MAX-DOAS retrievals, vertical smoothing difference errors reduce and the negative bias of -44% improves to -25%. **Figure 38** illustrates the difference in vertical sensitivity to the HCHO column, by showing typical vertical averaging kernels for TROPOMI, FTIR and MAX-DOAS retrievals.

7.3.3.3 OMI QA4ECV Data Record

The S5P TROPOMI L2_HCHO OFFL data (RPRO processor version 01.01.05 + OFFL processor version 01.01.05 to 01.01.07, covering one full year from the phase E2 and the operational phase) are compared with the OMI QA4ECV HCHO product (Nov.2004 - Dec.2018). The latter offers 15 years of afternoon observations with consistent algorithms, sharing the same auxiliary datasets (except for the cloud products).

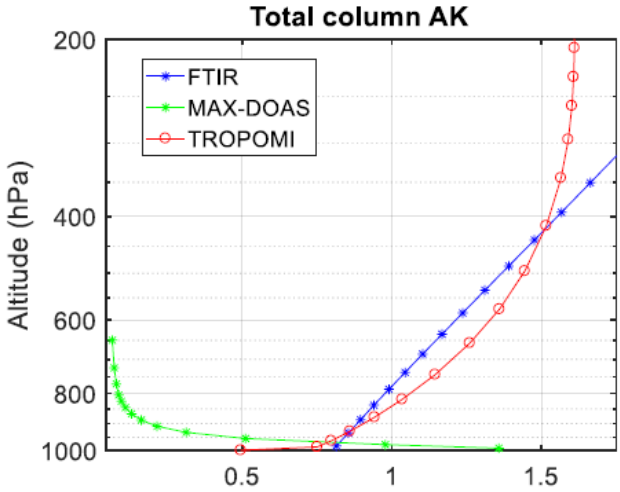


Figure 38: Typical total column averaging kernels for S5P TROPOMI and the two ground-based instrument types: FTIR (blue), MAX-DOAS (green), and TROPOMI (red). This illustrates the problem of the vertical smoothing difference error in these comparisons, as the instruments "see" different parts of the column.

Table 6 summarizes the correlation and slope of the QA4ECV OMI-TROPOMI comparison for a selection of regions. Two main effects are likely to cause remaining differences between the QA4ECV OMI and TROPOMI datasets. (1) The cloud products are from different algorithms, impacting the AMF mainly through cloud pressure differences. (2) For the background correction, different polynomial orders are used to compute the latitudinal dependency, causing a zonal dependency in the differences. Therefore, 3 comparisons are provided in order to quantify the 2 effects: VCD from OMI v1.1 and TROPOMI (different cloud products, different polynomials), clear VCD from OMI v1.1 and TROPOMI (no cloud correction, different polynomials), clear VCD from OMI v1.2 and TROPOMI (no cloud correction, same polynomials).

As can be seen in **Table 6**, the final bias between OMI and TROPOMI is below 15% in all 25 regions and below 10% in 21 regions. In mid-latitude regions, differences from 10% to 30% to could be attributed to the background correction polynomial. They have been reduced by aligning the OMI algorithm on TROPOMI (OMI v12). Differences due to the cloud correction can be observed in Tropical regions (up to 25%) and to a lesser extent in China. It is advised to use clear VCD when comparing satellite dataset using different cloud products.

Table 6 – Summary of the OMI-TROPOMI comparison statistics (correlation, slope) for a selection of regions. Numbers are given from phase E2 and forward. 3 Comparisons are provided: VCD from OMIv1.1 and TROPOMI, clear VCD from OMIv1.1 and TROPOMI, clear VCD from OMI v1.2 and TROPOMI, reflecting the effect of the cloud correction (VCD or clear VCD) and of the background correction (OMI v1.1 or OMI v1.2).

Region	Latitudes [degree]		Longitudes [degree]		VCD		VCD clear		Background Correction		Dispersion of the difference
	limits		limits		Correl.	Slope	Correl.	Slope	Correl.	Slope	[Pmolec.cm^2]
Northern Australia	-20.0	-12.5	125.0	145	0.79	1.08	0.72	0.97	0.69	0.97	1.38
Amazonia BB	-16.0	-6.0	-65.0	-45.0	0.86	1.18	0.87	0.99	0.92	0.94	1.30
Amazonia Bolivia	-14.0	-7.0	-66.0	-60.0	0.69	1.15	0.66	0.96	0.88	0.96	2.44
Southern Africa	-15.0	-5.0	10.0	30.0	0.52	1.20	0.70	1.01	0.91	0.97	1.10
Equatorial Pacific	-15.0	15.0	-180	-120	0	0.99	0	1.00	0.28	0.98	0.20
Amazonia	-10.0	5.0	-75.0	-50.0	0.85	1.19	0.84	0.98	0.92	0.94	0.91
Western Amazonia	-10.0	2.0	-75.0	-65.0	0.82	1.23	0.84	1.01	0.85	0.96	1.56
Equatorial Africa	-5.0	5.0	14.0	28.0	0.56	1.25	0.59	1.01	0.74	0.94	1.52
Indonesia	-5.0	5.0	98.0	118.0	0.31	1.08	0.35	1.03	0.52	0.97	0.67
Northern Africa	5.0	12.0	-14.0	12.0	0.57	1.10	0.55	0.94	0.84	0.91	1.00
South Asia	12.0	20.0	98.5	107.5	0.43	1.14	0.60	0.96	0.73	0.99	1.75
Guatemala	12.5	17.5	-95.0	-85.0	0.20	1.06	0.11	0.94	0.26	0.97	1.94
India	15.0	24.0	75.0	85.0	0.68	1.09	0.72	0.98	0.74	0.98	1.81
Mexico	15.0	20.0	-103	-88.0	0.41	1.09	0.42	0.97	0.55	0.96	1.58
Northern India	20.0	28.0	77.0	92.0	0.64	1.14	0.75	1.03	0.83	0.99	1.62
Southern China	24.0	33.0	108.0	121.0	0.56	1.26	0.60	1.08	0.62	1.00	2.67
Northern China	29.0	37.0	112.0	121.0	0.47	1.24	0.59	1.10	0.71	1.03	2.87
Southeastern US	30.0	40.0	-95.0	-75.0	0.86	1.30	0.91	1.12	0.92	1.05	1.74
Central US	30.0	44.0	-110	-95.0	0.80	1.33	0.80	1.19	0.84	1.10	1.79
Western US	34.0	44.0	-125	-110	0.76	1.23	0.72	1.17	0.84	1.10	1.78
Europe	35.00	50.00	0.0	40.0	0.69	1.32	0.68	1.26	0.89	1.13	1.01
Southeastern Canada	45.00	55.00	-90.0	-70.0	0.84	1.44	0.83	1.27	0.85	1.14	1.77
Northwestern Canada	60.00	70.00	-130	-95.0	0.37	0.63	0.33	0.54	0.51	0.99	1.71
Central Siberia	60.00	70.00	50.0	80.0	0.48	0.71	0.47	0.65	0.56	0.92	1.93

7.3.4 Dispersion

7.3.4.1 Fourier Transform Infrared Spectrometers

Using correlative measurements at the 25 NDACC FTIR stations, the median absolute deviation (MAD) is close to the mission requirement of 12 Pmolec/cm^2. While the MPC Automated Validation Server performs comparisons using single TROPOMI pixels, further investigation in the NIDFORVAL

AO project use an average of about 30 TROPOMI pixels (20 km around the station). The MAD for all stations is of 2.4 Pmolec/cm^2 while the requirements is 12/sqrt(n pixels)= 2.1 Pmolec/cm^2. However, it is more relevant to compare the MAD obtained at the clean sites only to evaluate the TROPOMI precision because MAD is less sensitive to additional co-location mismatch errors in regions far from emissions. In clean conditions, the TROPOMI precision is found to be much better that the pre-launch mission requirements: 1.3 Pmolec/cm^2, corresponding to a single pixel precision of 7 Pmolec/cm^2. Details can be found in Vigouroux et al. (2020).

7.3.4.2 MAX-DOAS UV-visible Spectrometers

The dispersion (standard deviation) of the difference of S5P with respect to MAX-DOAS is 8 Pmolec/cm^2 at Cabauw and 10 Pmolec/cm^2 at De Bilt. This is within the mission requirement of precision of 12 Pmolec/cm^2.

7.3.4.3 OMI QA4ECV Data Record

As reported in **Table 6** the dispersion of the daily difference between OMI and TROPOMI is generally ranging from 1 to 2 Pmolec/cm^2, with the exception of Northern China (4 Pmolec/cm^2). Low dispersion is related to the large number of observations included in the averages. The standard deviation of individual OMI and TROPOMI observations is respectively about 7 and 4 Pmolec/cm^2 in remote regions with no local emissions. 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

Day to day correlation between OMI and TROPOMI is very high above emission regions. Overall, the short term variability seen in the MAXDOAS measurements is nicely reproduced by S5P TROPOMI. **Figure 39** presents seasonal cycle plots of S5P HCHO and MAXDOAS tropospheric VCD. In July-August, MAX-DOAS is higher than S5P HCHO. The seasonal variability captured by TROPOMI is similar to the one reported by the FTIR and MAX-DOAS instruments, with correlations of 0.92 and 0.83, respectively.

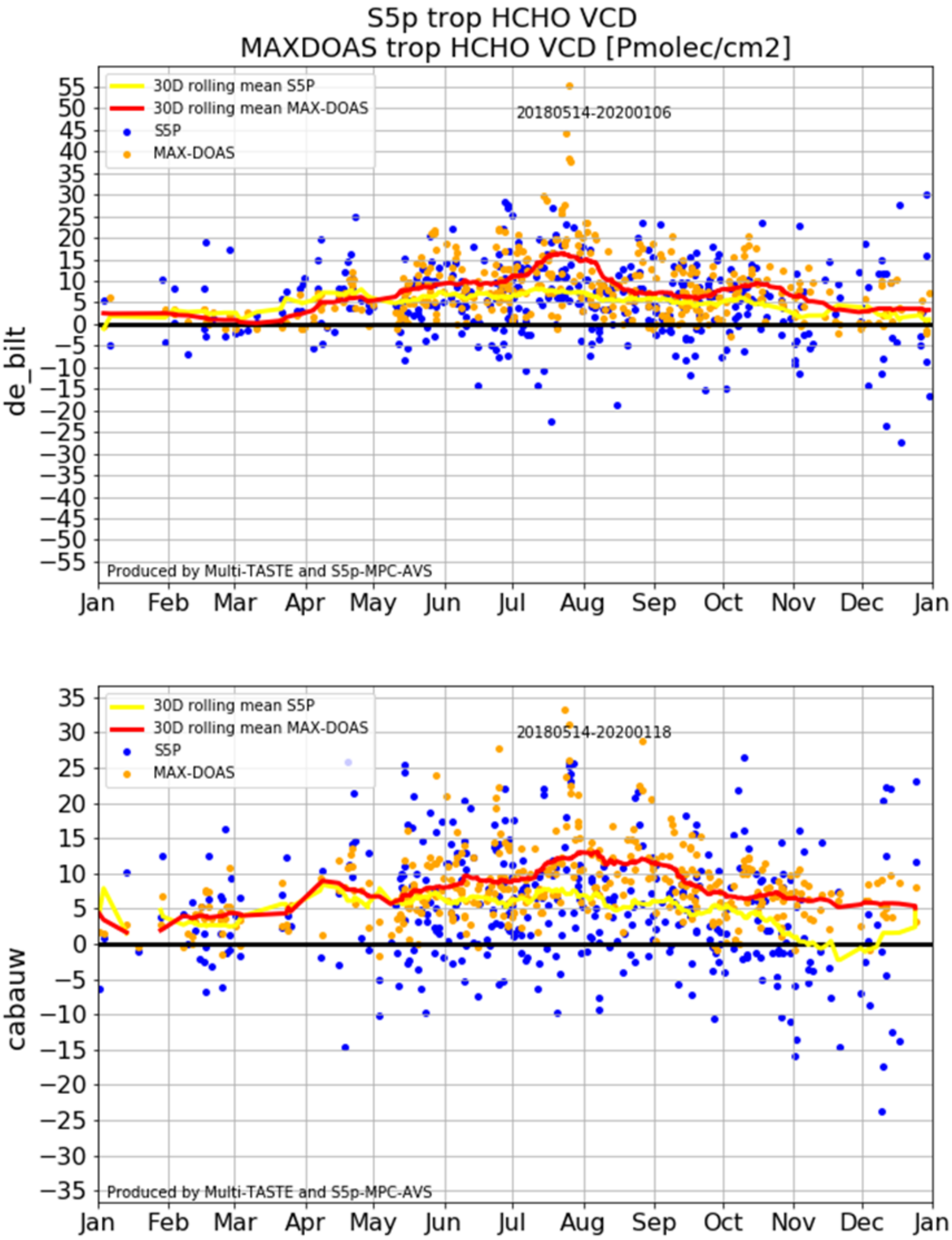


Figure 39: Seasonal cycle plots of S5P HCHO tropospheric column and MAXDOAS measurements at De Bilt and Cabauw. The 30-day rolling mean (cyclic boundary condition) is also indicated.

7.3.7 Geographical patterns

The S5P_L2_HCHO data are seasonally averaged for spring (March-May 2018) and summer (June-August 2018) and compared to OMI (González Abad et al., 2015, 2016) and GOME-2. The comparison results are shown in **Figure 40** and **Figure 41**. The results show similar spatial patterns of HCHO columns for the three satellites. Compared to OMI, TROPOMI observations show higher HCHO columns over India and the Sahara Desert. GOME-2 reports similar HCHO column values as TROPOMI in the same regions.

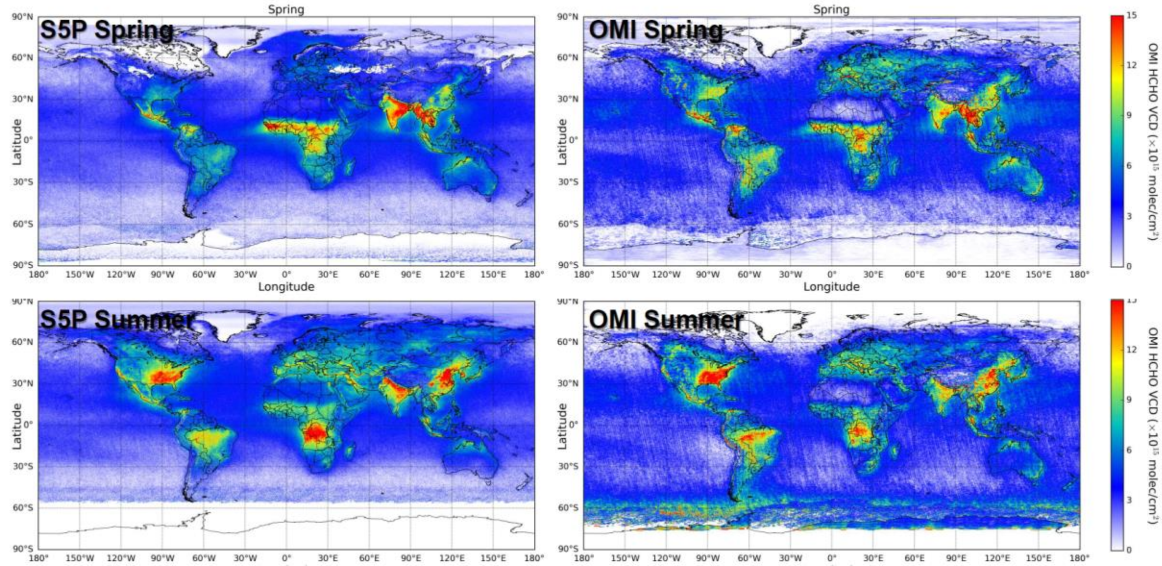


Figure 40: Seasonal average of HCHO total column data for S5P TROPOMI (left) and for OMI (right).

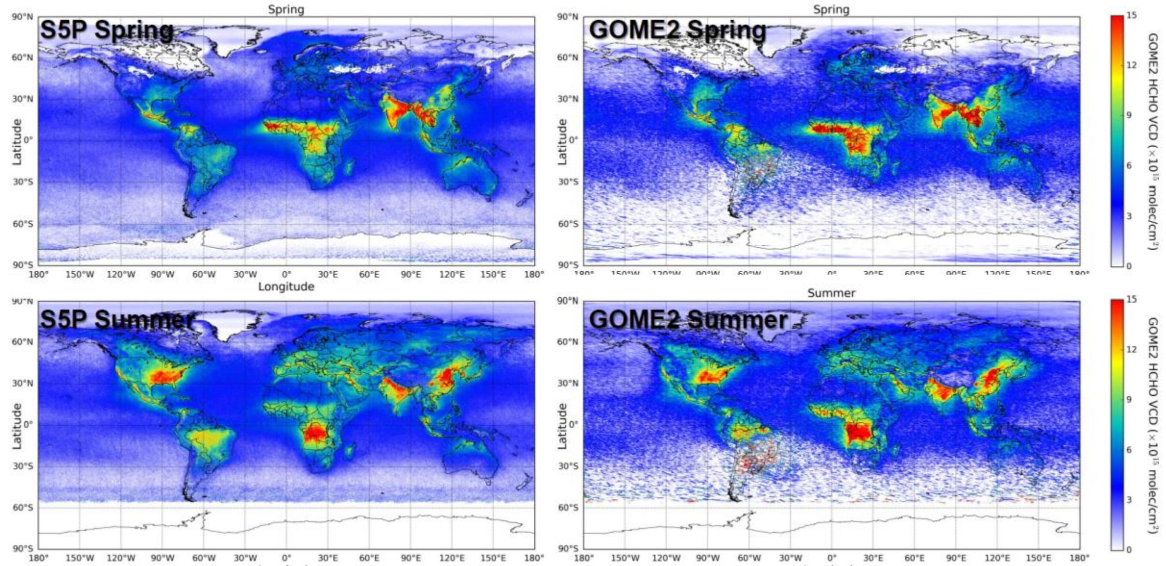


Figure 41: Seasonal average of HCHO total column data for S5P TROPOMI (left) and MetOp-B GOME 2 (right).

7.3.8 Other features

None to report.

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 (processor version 01.01.02 to 01.01.07) and L2_HCHO OFFL (RPRO processor version 01.01.05 + OFFL processor version 01.01.05 to 01.01.07), 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.**

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 validation server was consulted on 2020-02-28.

Cabauw: 287 common co-locations Orbits range from 4839 (2018-09-19) to 11735 (2020-01-01)			
	NRTI vs OFFL	NRTI vs MXD	OFFL vs MXD
Mean(diff)±sem*	-0.14	-3.84±0.48	-3.70±0.48
Median(diff)*	-0.04	-4.84	-4.86
Std(diff)*	1.6	8.1	8.1
1/2 IP68(diff)*	1.3	7.1	7.4
Pearson R	0.98	0.37	0.37
Slope	0.98	0.57	0.57
De Bilt: 268 common co-locations Orbits range from 4839 (2018-09-19) to 11564 (2020-01-06)			
	NRTI vs OFFL	NRTI vs MXD	OFFL vs MXD
Mean(diff)±sem*	-0.04	-1.08±0.55	- -1.04±0.56
Median(diff)*	0.06	-1.22	-0.90
Std(diff)*	1.9	9.1	9.2
1/2 IP68(diff)*	1.4	7.9	7.8
Pearson R	0.98	0.22	0.22
Slope	0.96	0.34	0.34

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).

7.4.2 Dispersion

Standard deviation and the ½ 68% interpercentile (1/2 IP68) of the NRTI-OFFL differences are much smaller than that between either NRTI and MAX-DOAS or OFFL and 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.

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.

8.2 Validation approach

8.2.1 Ground-based networks

Boundary layer pollution (SO_2 total)

S5P TROPOMI L2_SO2 sulphur dioxide column data are compared to ground-based MAX-DOAS UV-visible observations. However, currently the number of available stations in strongly polluted regions is very rare. Outside strongly polluted regions, the SO_2 column is below the detection limit of both the MAX-DOAS and satellite measurements. For the validation of the S5P TROPOMI L2_SO2 sulphur dioxide column data MAX-DOAS measurements at Xianghe (China), Greater Noida (India), and Basra (Iraq) were used so far.

Volcanic plumes (SO2 enhanced)

S5P TROPOMI L2_SO2 sulphur dioxide column data are compared to MAX-DOAS UV-visible measurements collected from the Network for Observation of Volcanic and Atmospheric Change (NOVAC) [ER_NOVAC]. Because of the strong SO_2 concentration gradients in volcanic plumes, the comparison is not performed using the SO_2 columns but rather using the derived SO_2 fluxes.

8.2.2 Satellites

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

8.2.3 Field campaigns and modelling support

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

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

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

8.3 Validation of L2_SO2 NRTI

8.3.1 Recommendations for data usage followed

The quality of the observations depends on many factors which are taken into account in the definition of the `qa_value`. While it is a handy way of filtering observations of less quality, the “quality assurance value” should also be considered with caution, as it is a compromise to take into account several aspects, such as: processing errors, presence of clouds or snow/ice, observations affected by sun glint, South Atlantic Anomaly, possible contamination by volcanic SO_2, absence of background correction, and important variables out of range (importantly the AMF).

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

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

8.3.2 Status of validation

So far the validation of the S5P TROPOMI L2_SO2 sulphur dioxide column data is mainly based on satellite to satellite comparisons, 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 SO_2 columns were mostly close to or below the detection limit of S5P TROPOMI.

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

Outside strongly polluted regions and volcanic plumes, the atmospheric SO_2 SCDs were found to be consistent with the assumption of zero within the measurement uncertainties.

From these comparisons (details are shown below) the following conclusions are drawn:

- over polluted regions the requirements are fulfilled;
- over volcanic plumes the bias requirement is fulfilled, 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 SO_2 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 SO_2 SCDs (and their errors and standard deviations) it is concluded that the requirements are fulfilled. The bias and spread are typically below 0.2 DU.

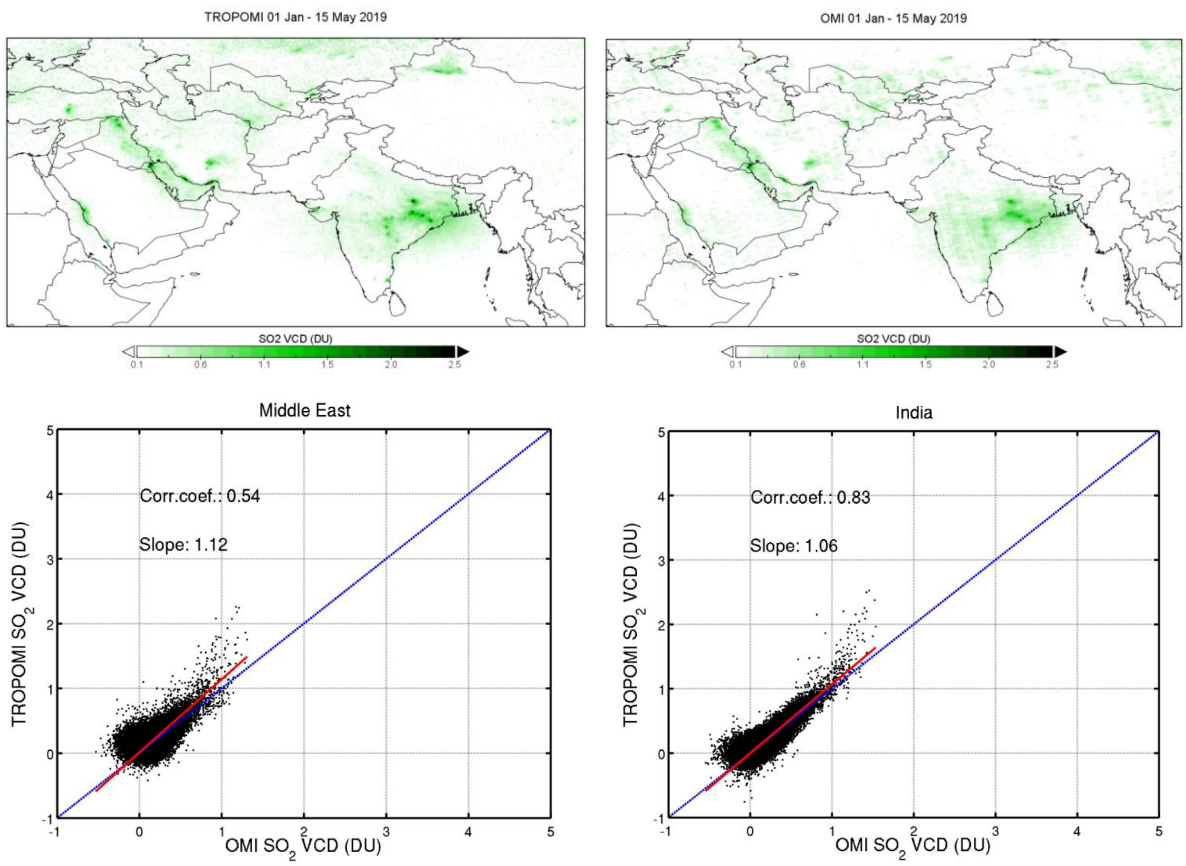


Figure 42: Top: Comparison of the average distribution (01 Jan 2019 – 15 May 2019) of the SO_2 VCDs derived from TROPOMI and OMI over regions with strong air pollution. Both data sets show very good agreement. Bottom:Correlation plots TROPOMI versus OMI over the Middle East and India. Note that a fixed AMF of 0.4 was used for both retrievals to exclude the effect of different profile assumptions. Courtesy of Nicolas Theys, BIRA-IASB.

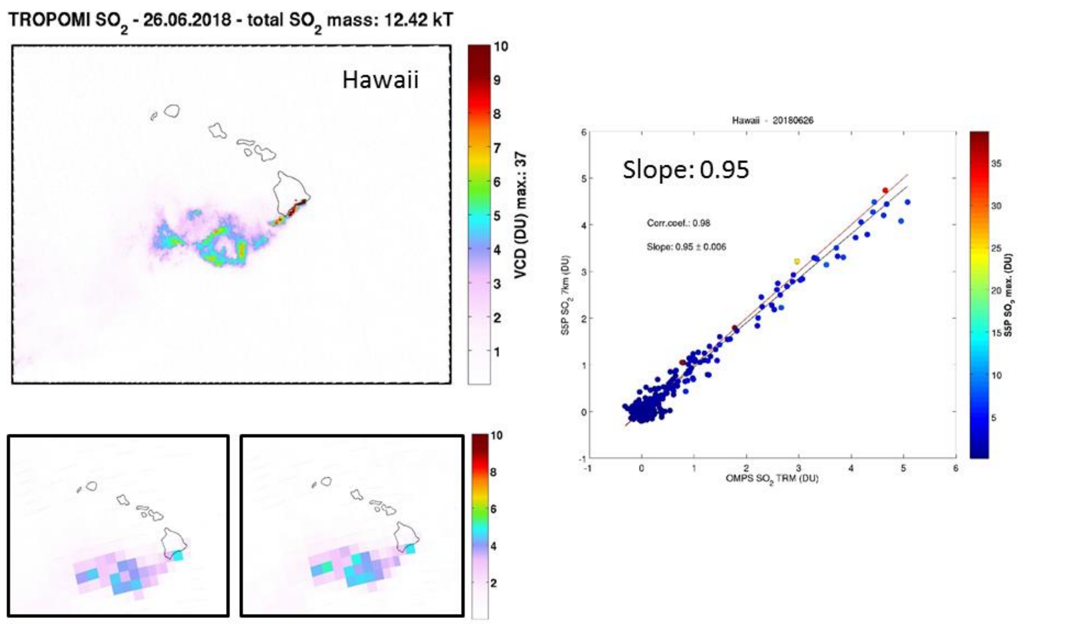


Figure 43: Comparison of TROPOMI and OMPS measurements of the volcanic plume of Kilauea on 26 June 2018. The large figure shows the original TROPOMI data. The two small figures show the spatially degraded TROPOMI data and the OMPS data. The figure right shows the correlation plot of the degraded TROPOMI data versus the collocated OMPS data. Courtesy of C. Li and N. Krotkov, NASA/GSFC.

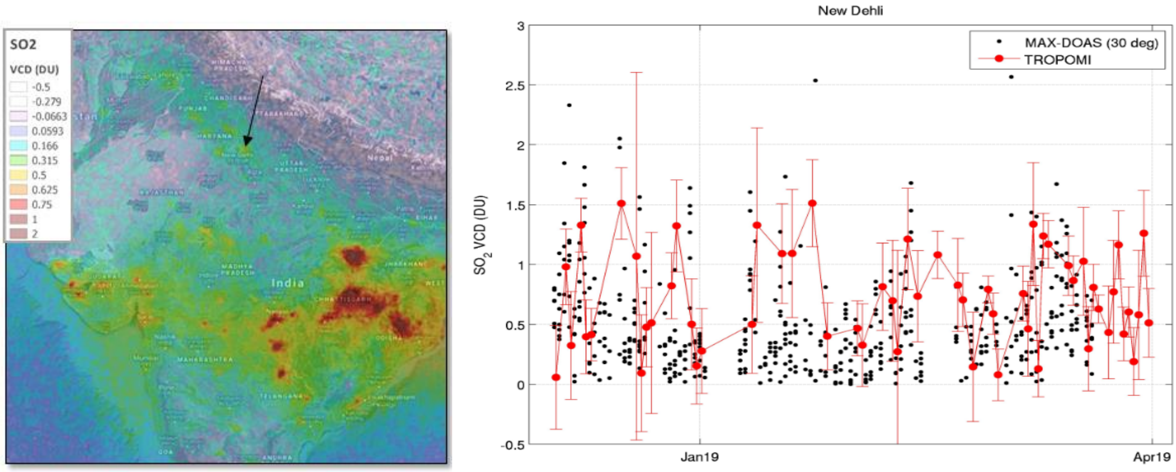


Figure 44: Comparison of TROPOMI SO_2 VCDs to MAX-DOAS measurements at Greater Noida (close to New Delhi, India). The following selection criteria were applied: distance < 15km, CF<0.2, AMF>0.2, MAX-DOAS +/- 1h around S5P overpass. Courtesy of M. Sharma (Sharda University, India) S. Donner, S. Dörner, T. Wagner (MPIC), N. Theys (BIRA-IASB).

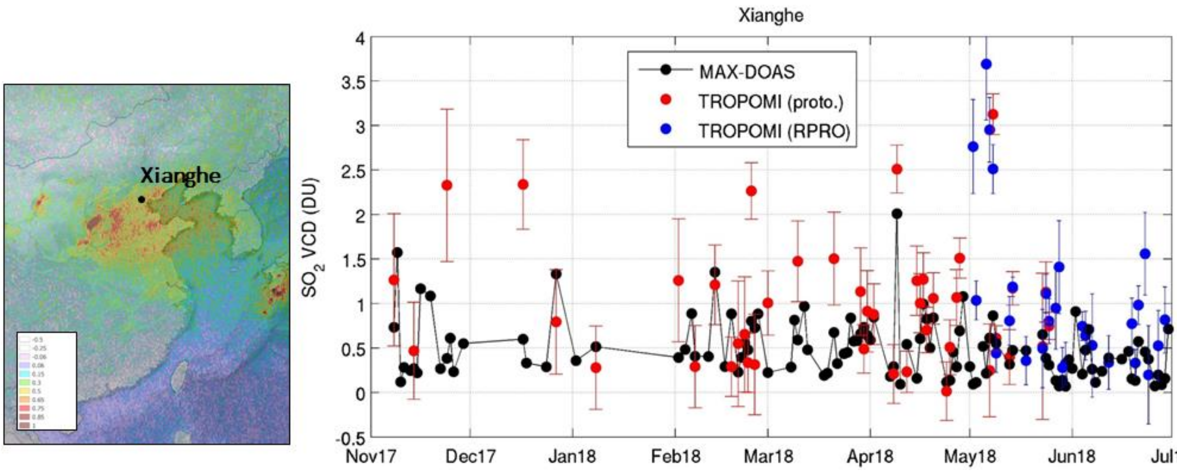


Figure 45: Comparison of TROPOMI SO_2 VCDs to MAX-DOAS measurements (daily means) at Xianghe (China). The following selection criteria were applied: distance < 15km, CF<0.2, AMF>0.2, number of observations >10. Courtesy of N. Theys (BIRA-IASB).

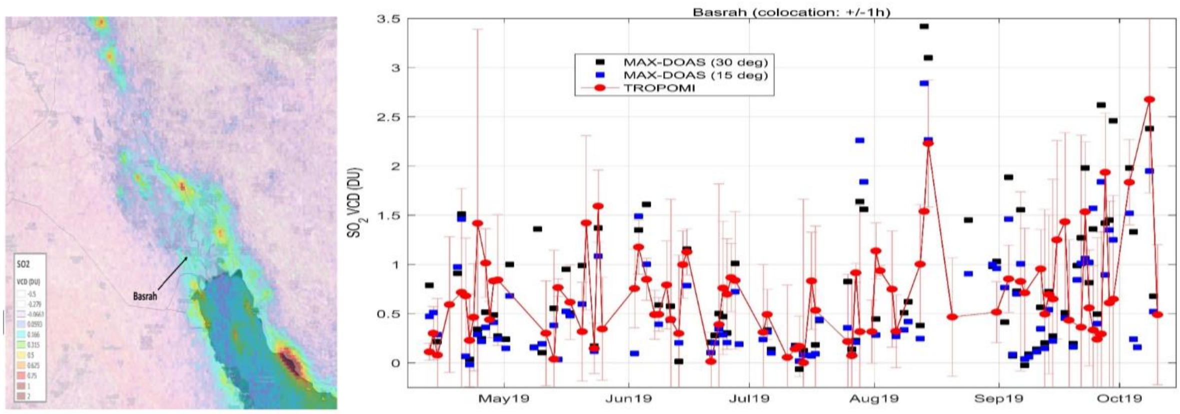


Figure 46: Comparison of TROPOMI SO_2 VCDs to MAX-DOAS measurements (daily means) at Basra (Iraq). The following selection criteria were applied: distance < 25km, CF<0.2, AMF>0.2, time window +/- 1h around overpass. Courtesy of N. Theys (BIRA-IASB), data provided by Nayyef Almaliki, Mustafa Aldossary, Ali Almasoudii, Sebastian Donner, Steffen Dörner, Thomas Wagner.

8.3.3 Bias

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

8.3.4 Dispersion

The dispersion is well within requirements for observations boundary pollution. For observations of strong volcanic plumes the dispersion 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 SO_2 SCDs it is estimated that the dispersion is within 0.2 DU.

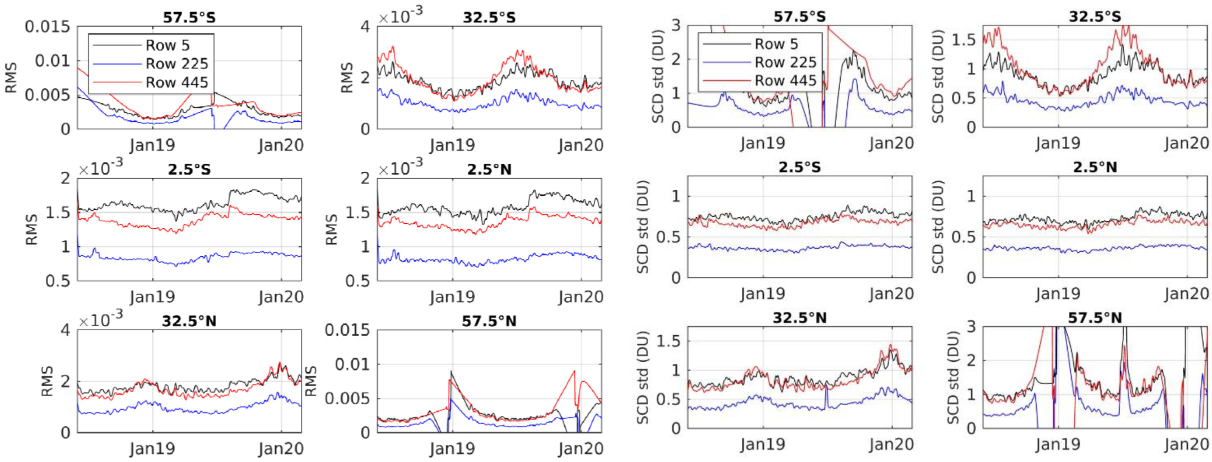


Figure 47: Temporal evolution of the measurement error (left) and the standard deviation (right) for selected 5° latitude bands and 3 detector rows from December 2018 to February 2020. Good qualitative agreement between both quantities is found indicating that the random uncertainty is well characterized by the measurement error. Larger errors (and standard deviations are found at the edges of the detector and towards high latitudes. Courtesy of N. Theys (BIRA-IASB).

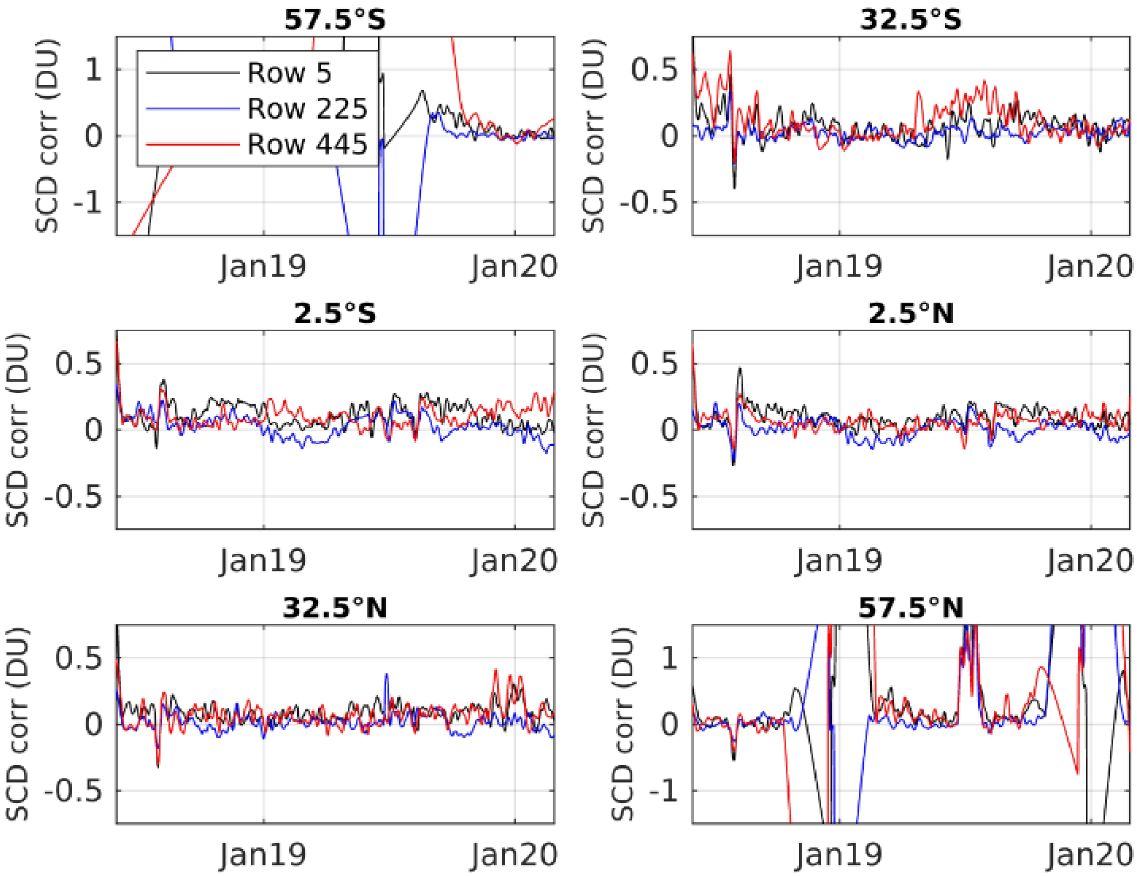


Figure 48: Temporal evolution of the averaged SO_2 SCD for selected 5° latitude bands and 3 detector rows from December 2018 to February 2020. The values are close to zero and show relatively small day to day variations. The larger variations in August are caused by strong volcanic eruptions. Courtesy of N. Theys (BIRA-IASB).

8.3.5 Dependence on influence quantities

Slightly larger bias and dispersion are found towards higher SZA.

8.3.6 Short term variability

The short term variability can be estimated from the time series of averaged SO_2 SCDs (outside periods with strong volcanic eruptions). It is estimated to below about 0.1 DU.

8.3.7 Geographical patterns

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

8.3.8 Other features

None to report.

8.4 Equivalence of L2_SO2 NRTI and OFFL products

The NRT and offline SO_2 products are very similar, as illustrated by the comparison of the SO_2 SCDs of both data versions hereafter. Thus, the validation activities performed for the OFFL data product (see above) are also representative for the NRTI data product.

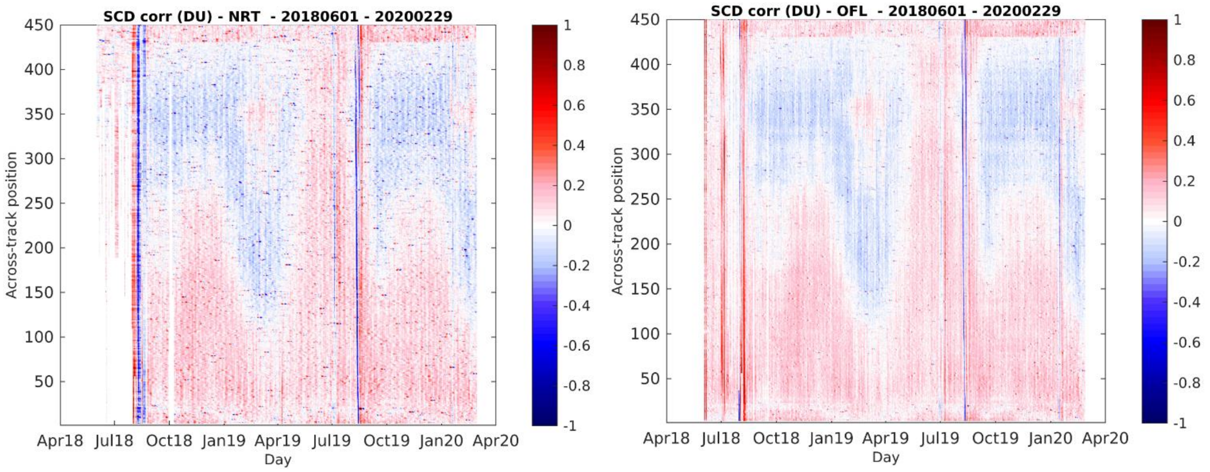


Figure 49: Comparison of the NRT (left) and offline (right) SO_2 data products. Shown are the time series of background corrected SO_2 SCDs for all 450 detector rows from June 2018 to February 2020. The short time features (vertical colored lines) are caused by individual strong volcanic eruptions. Courtesy 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 both 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 50** 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%.

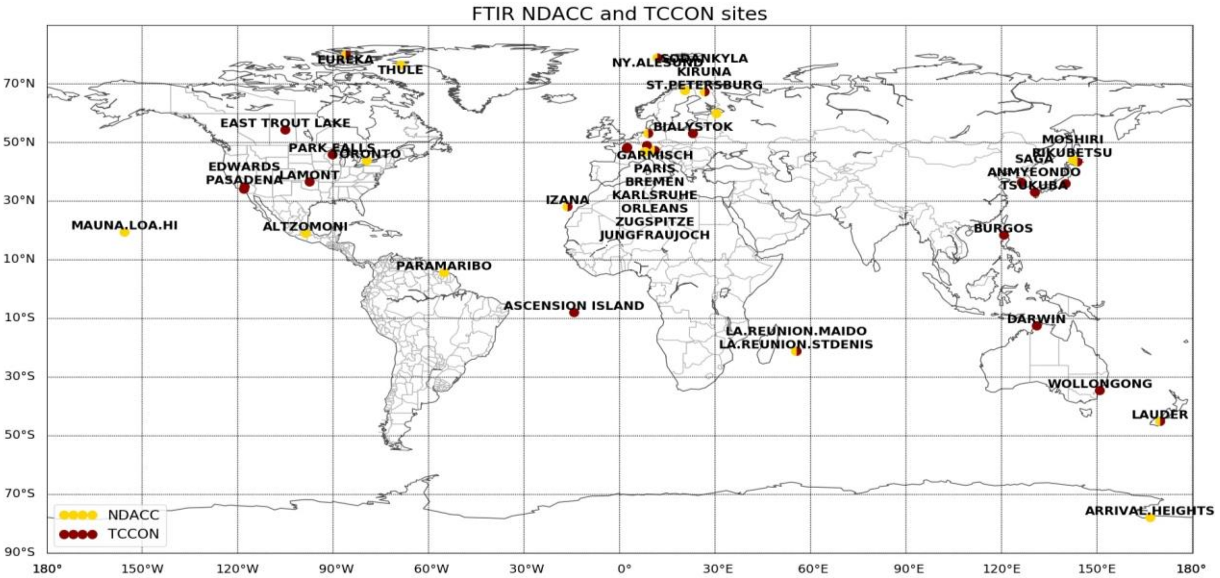


Figure 50: Geographical distribution of NDACC and TCCON FTIR stations measuring atmospheric carbon monoxide column data. Some sites contribute to both networks.

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 at http://mpc-vdaf.tropomi.eu.

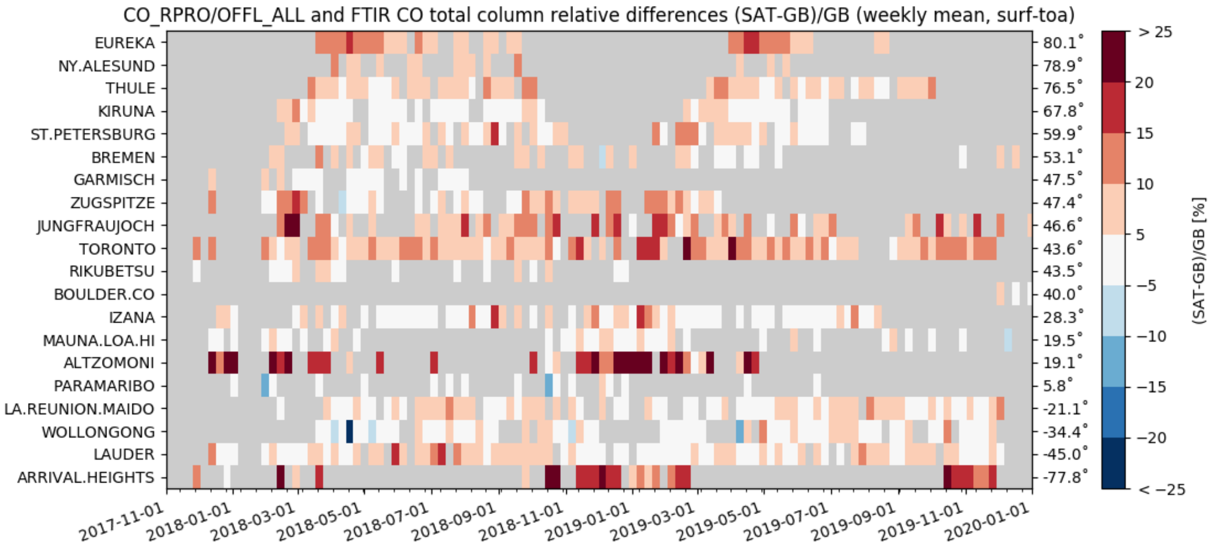
Current conclusions are based on the amount of reference measurements available at the time of this analysis, yielding comparison pairs from November 2017 through December 2019. Routine validation is done using the Automated Validation Server of the MPC VDAF, the CO validation system operated at BIRA-IASB, 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 6h 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. Given the limited number of co-locations with TCCON and NDACC measurements, this needs to be confirmed in future reports.

9.3.3 Bias

The systematic difference between S5P L2_CO daily mean data and correlative ground-based measurements is on an average 6.5% with respect to NDACC data and 8.8% 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 52** and **Figure 53** show the biases over the time period Nov 2017–December 2019 sorted by latitude. It seems that there is no latitudinal dependence of the bias. **Figure 51** does not show any significant degradation in bias with time (note that the longer time period covers different processor versions). **Figure 51** also shows a slight increase of bias during local winter, but due to the limited time period it is too early to make conclusions on seasonal cycle of the bias.



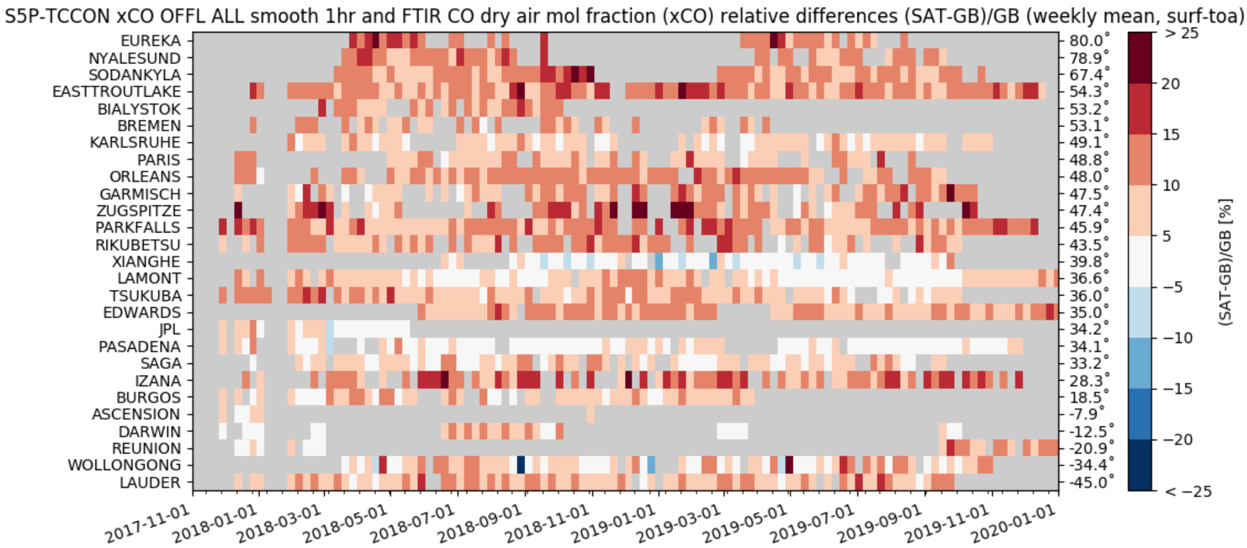


Figure 51: Relative bias between S5P L2_CO OFFL and ground-based CO column data at NDACC (top) and TCCON (bottom) FTIR stations. Over the Nov. 2017 – December 2019 time period the plots do not show a clear meridian dependence or temporal change in the weekly averaged biases.

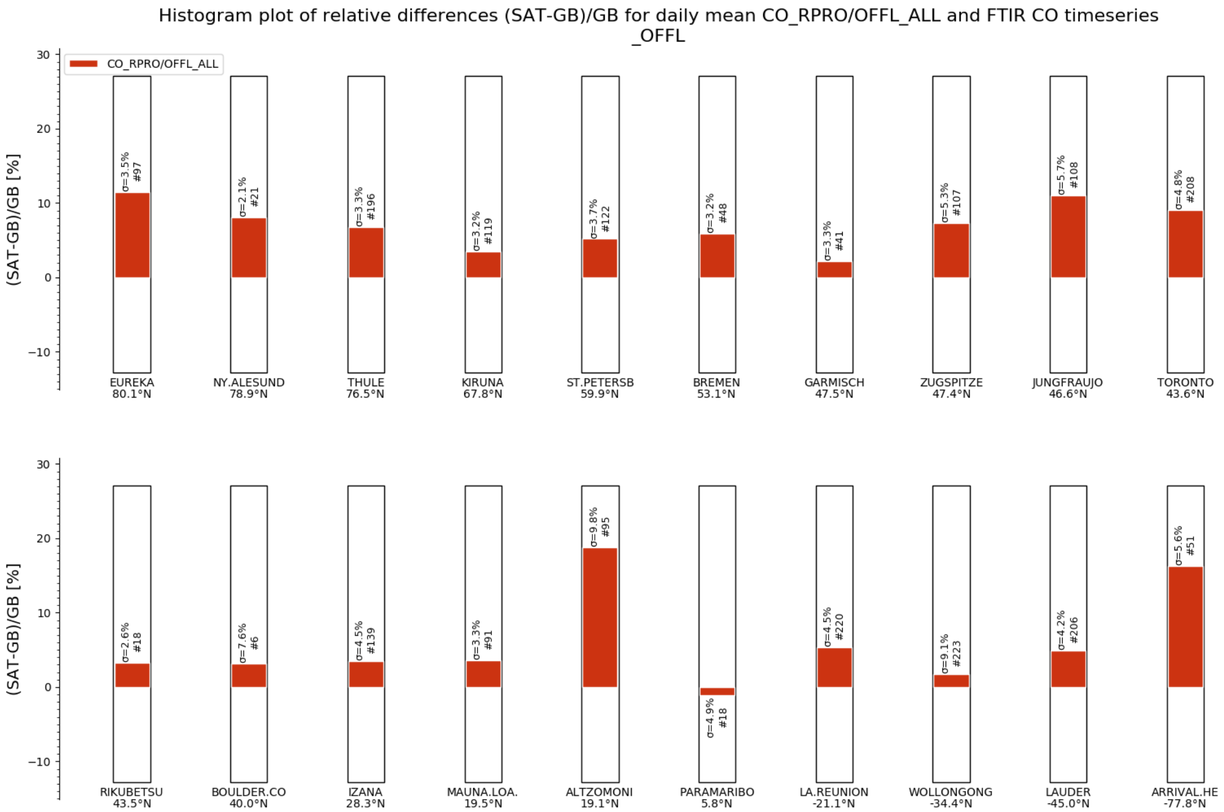


Figure 52: Bar chart of the relative mean difference between S5P L2_CO OFFL and FTIR CO column data at 20 NDACC sites, for all data within the time range from December 2017 till December 2019 showing RPRO/OFFL. The sites are sorted with decreasing latitude. All biases are below 15% except at Altzomoni, 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.

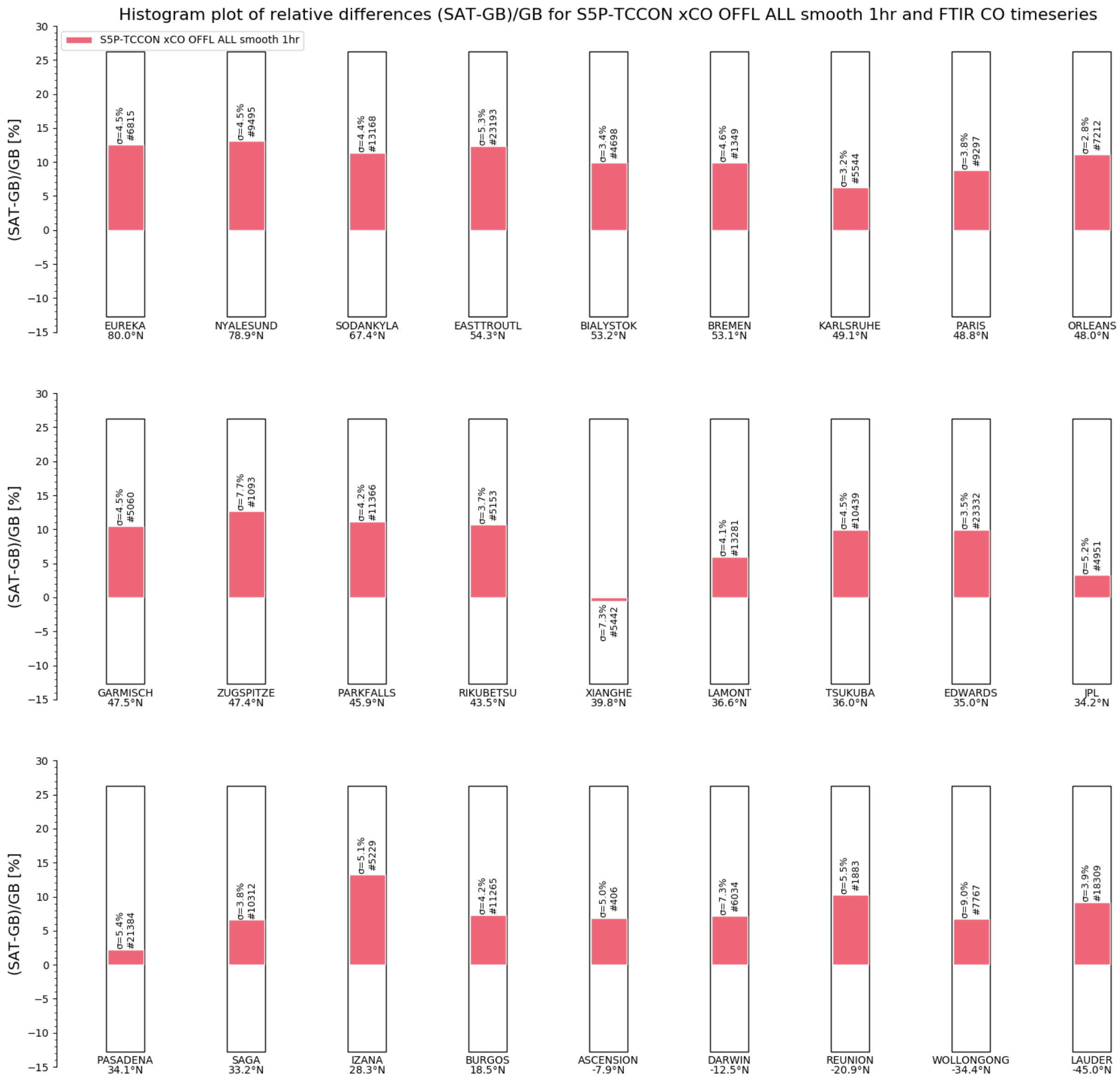


Figure 53: Bar chart of relative mean difference between S5P L2_CO OFFL and FTIR CO column data at 27 TCCON sites for all data within the time range Nov 2017 till December 2019. The sites are sorted with decreasing latitude. The majority of the OFFL biases are below 11% except in the Arctic and Izana (mountain station) where the bias is slightly above 11%. 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 52** and **Figure 53.** 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. A more precise estimate will be made when more measurement data is available.

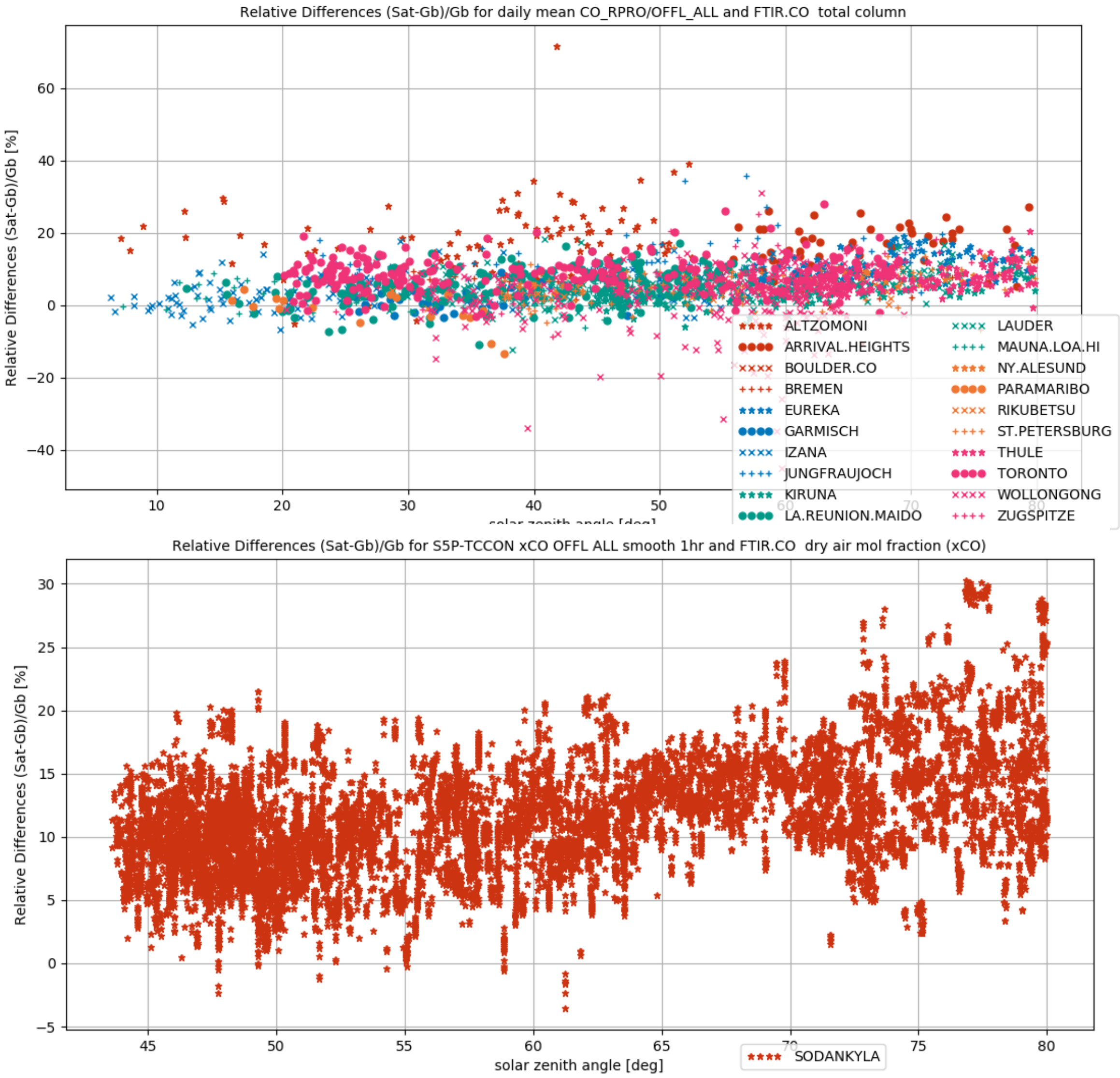


Figure 54: Relative difference (daily mean) between S5P L2_CO RPRO/OFFL and NDACC (top) or TCCON Sodankylä (bottom) carbon monoxide total column as a function of the TROPOMI solar zenith angle, in the 'all' case.

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 55**).

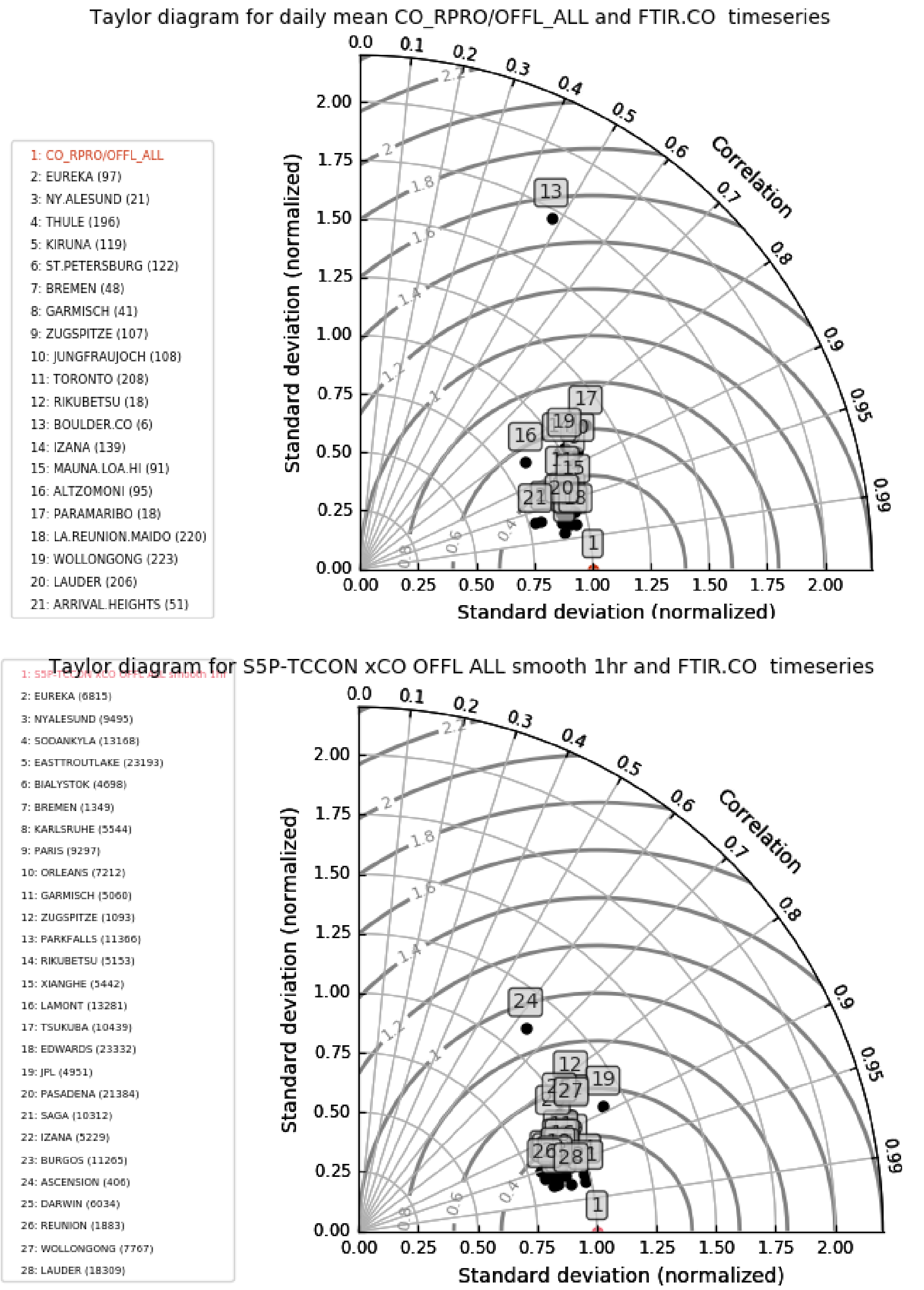


Figure 55: Taylor diagrams for daily mean differences between S5P L2_CO OFFL and ground-based networks CO data: NDACC (top) and TCCON (bottom) for our all case of pixel selection criteria (Boulder and Ascension cover a limited time period)

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 56** 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 57**.

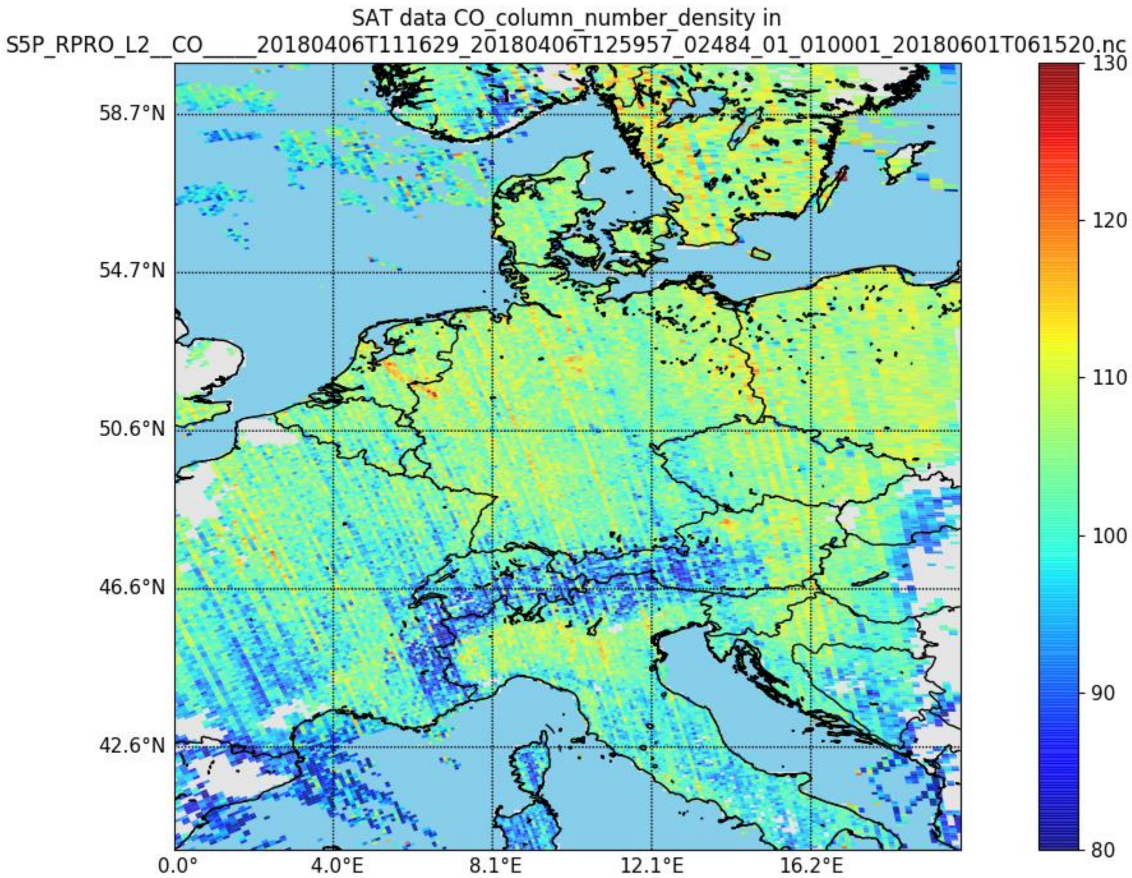


Figure 56: Example of stripe patterns in L2_CO OFFL CO column data along a S5P orbit above Europe.

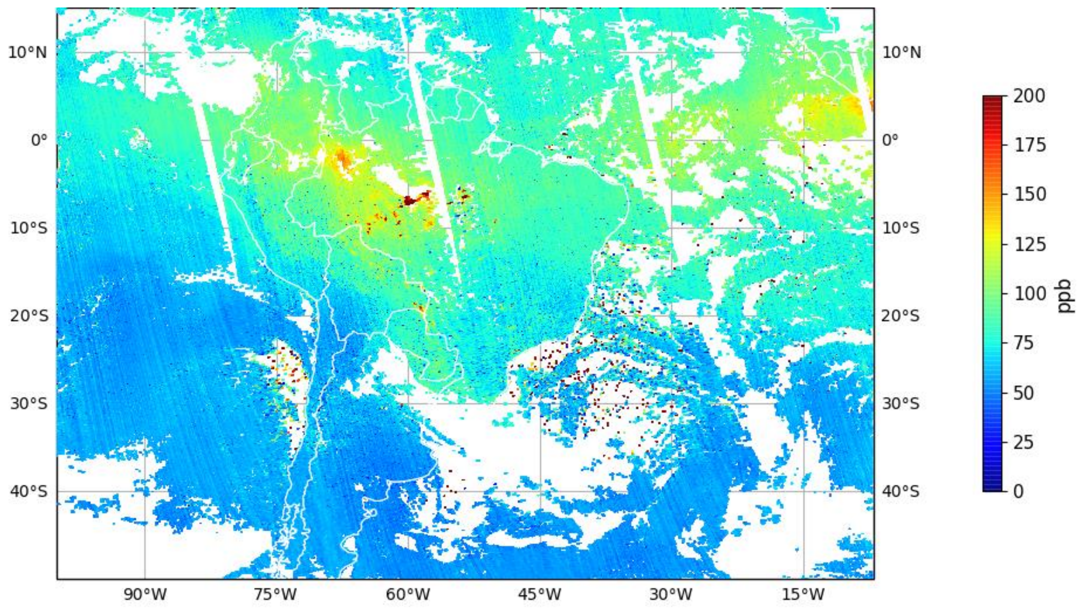


Figure 57: S5p OFFL xCO pixels measured on August 1, 2019, over South America and the Atlantic Ocean. Outlying pixels occurs (including negative values) in the South Atlantic Anomaly.

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

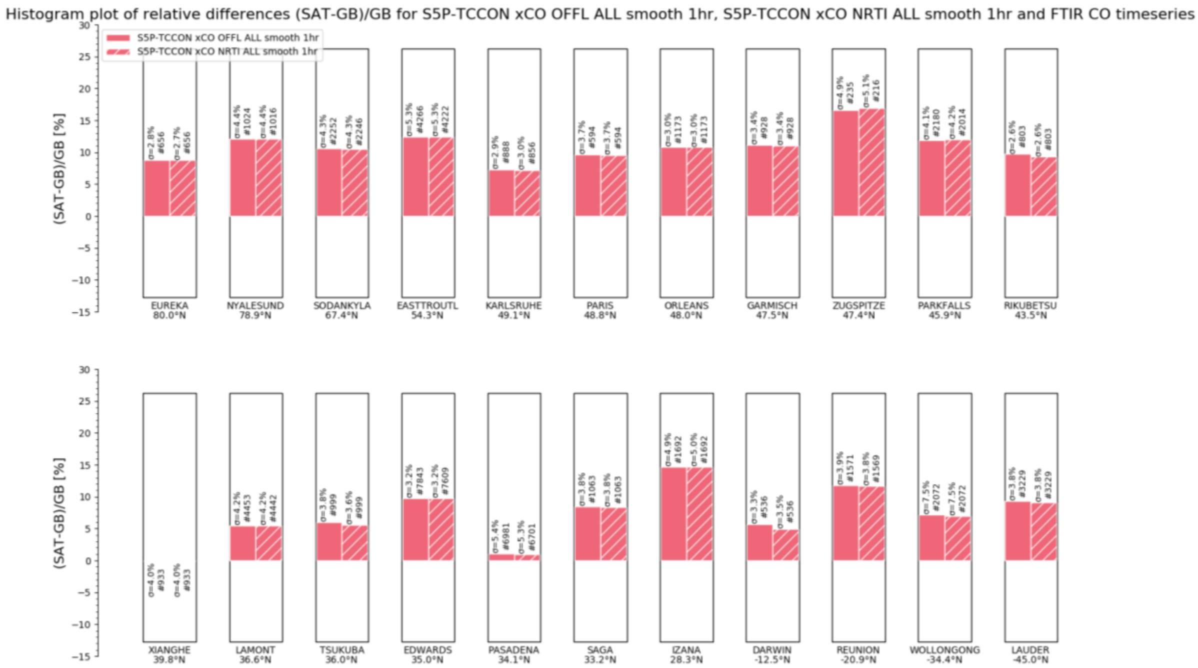


Figure 58: Comparison of relative biases against TCCON CO column data for the S5P L2_CO OFFL and NRTI data versions, from July 3 2019 through December 2019. The quality of both data sets is similar. Over this period the OFFL processor has produced data with a relative mean bias of 9.1% ± 4.06% and a correlation coefficient 0.91 with respect to TCCON data; the NRTI processor has produced data with a relative mean bias 9.0% ± 4.07% and a correlation 0.90.

10 Validation Results: L2_CH4

10.1 L2_CH4 products and requirements

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

10.2 Validation approach

10.2.1 Ground-based networks

S5P TROPOMI L2_CH4 methane column data are routinely compared to reference measurements obtained from FTIR spectrometers performing network operation in the context of the Network for the Detection of Atmospheric Composition Change (NDACC, http://ndacc.org) and the Total Carbon Column Observing Network (TCCON, https://tccondata.org). **Figure 50** displays the geographical distribution of the NDACC and TCCON sites. Near-infrared TCCON measurements provide calibrated methane column averaged (xCH_4) data with typical uncertainty values of 0.5% for the precision and 0.2% for the accuracy. Solar infrared NDACC measurements provide CH_4 total column data with a lower accuracy (typically 3%) and precision (1.5%). The required accuracy (<1.5%) and precision (<1%) for S5P implies that we mainly focus on the validation with TCCON measurements.

10.2.2 Satellites

None for this report.

10.2.3 Field campaigns and modelling support

None for this report.

10.3 Validation of L2_CH4 OFFL

10.3.1 Recommendations for data usage followed

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

The S5P L2 data contains two xCH_4 column values: the standard retrieved product and a bias corrected product. Both products are validated separately, but only the bias corrected is mentioned in the quality indicators in **Table 2**.

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

10.3.2 Status of validation

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

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

Current conclusions are based on the S5P and reference measurements available at the time of this analysis, which yield comparison pairs from January 2018 through December 2019. Routine validation is done using the Automated Validation Server of the MPC VDAF, the CH_4 validation system operated at BIRA-IASB, and the HARP toolset and shows an up-to-date comparison.

Since August 6 2019 (orbit 9388) S5P measures with increased spatial resolution from 7km to 5.5km along track. This change in operations did not change the performance of the methane OFFL product. Given the limited number of co-locations with TCCON and NDACC measurements, this needs to be confirmed in future reports.

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 -0.27% (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.

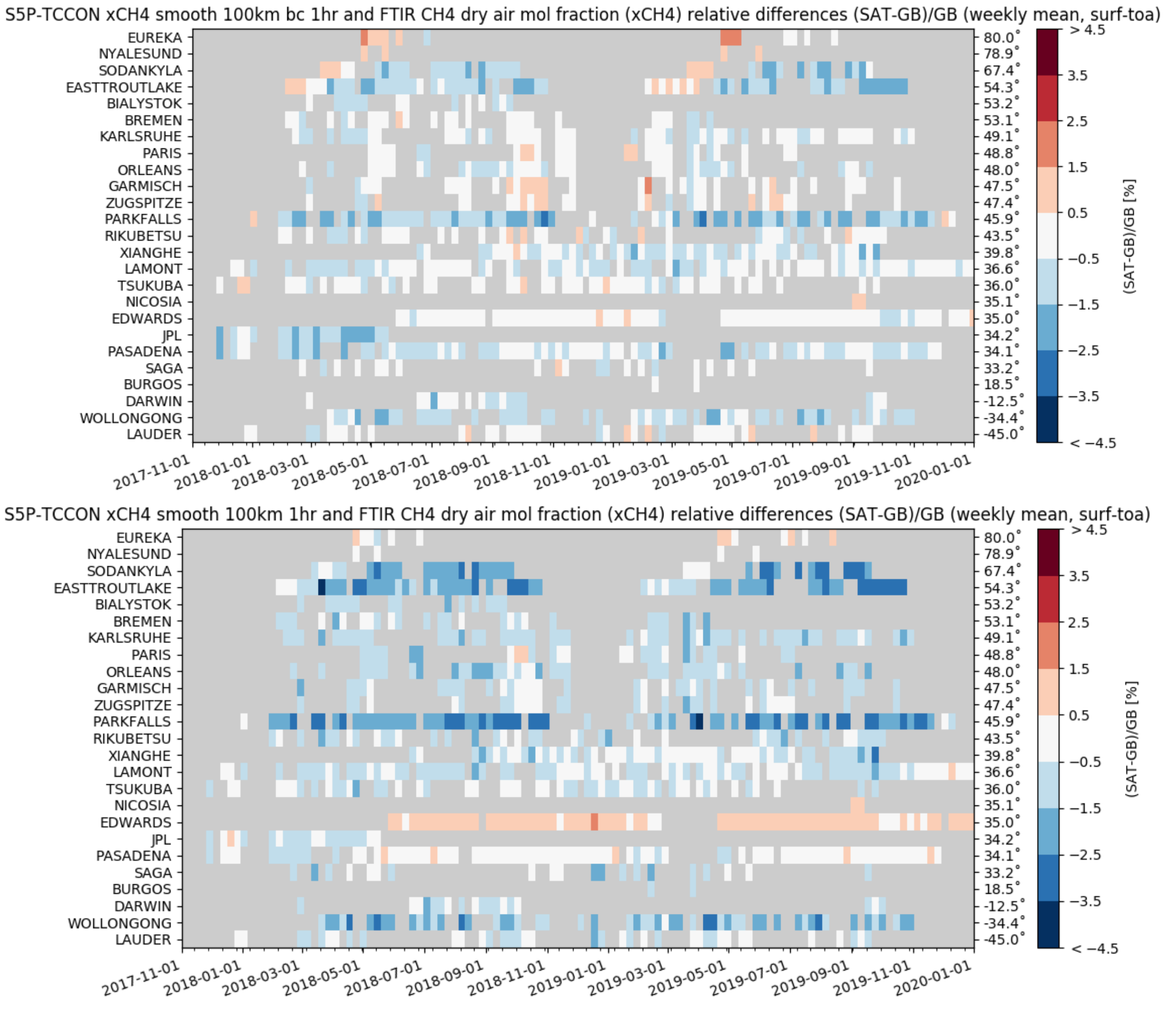


Figure 59: Mosaic plots of relative biases between S5P L2_CH4 RPRO+OFFL and ground-based CH_4 TCCON column data for the bias corrected (top) and standard (bottom) methane products. Over the November 2017 – December 2019 time period the plots do not show a clear meridian dependence or temporal change in the weekly averaged biases.

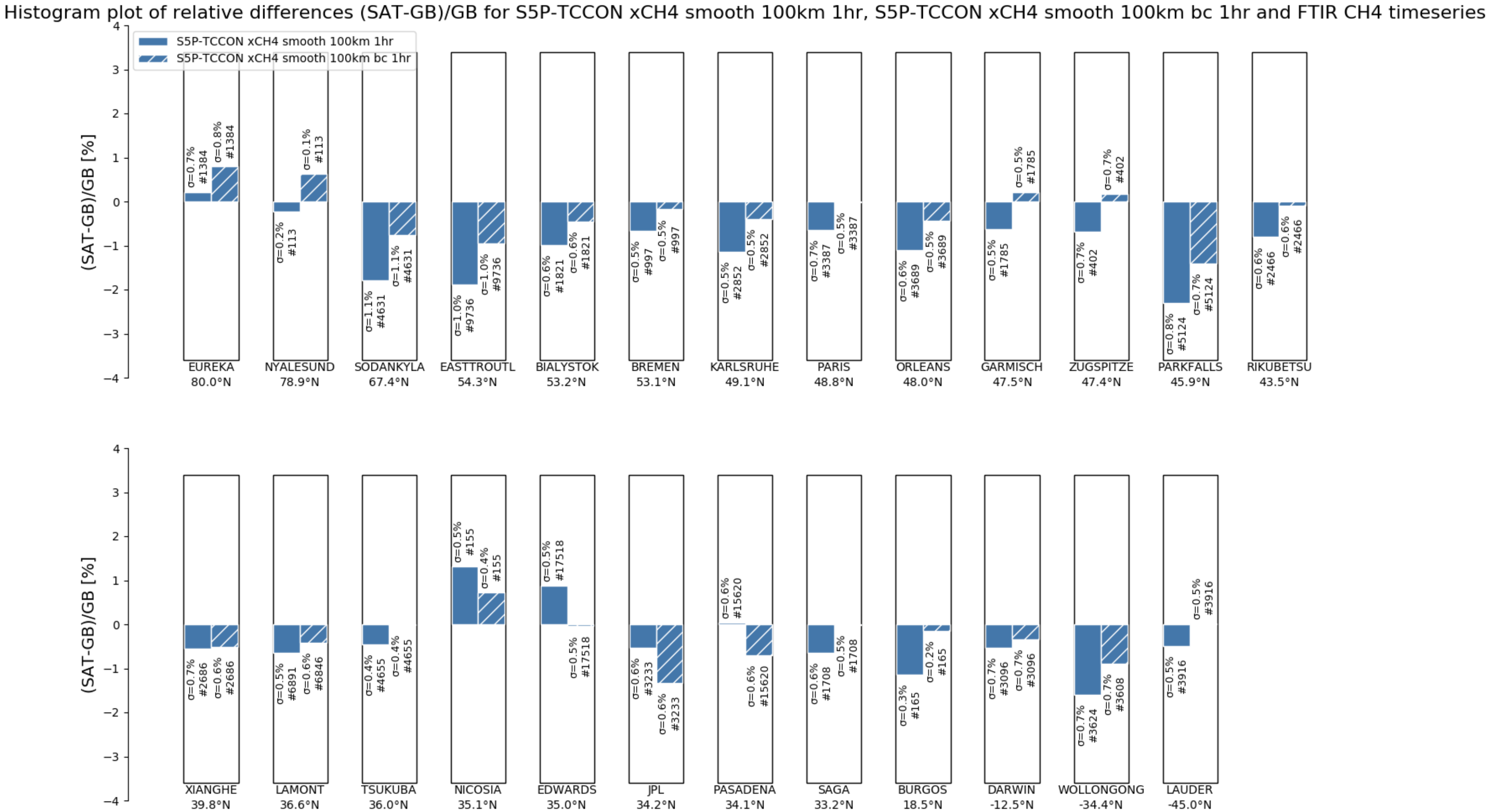


Figure 60: chart of relative mean difference between S5P L2_CH4 and FTIR CH_4 column data at 25 TCCON sites within the time range November 2017 till December 2019. The sites are sorted with decreasing latitude. The relative mean difference of the standard xCH4 product slightly exceeds the mission requirements (bias below 1.5%) only at a few TCCON sites (i.e. Sodankylä, East Trout Lake, Park Falls and Wollongong). However, it never exceeds the mission requirements for the bias corrected product.

	S5P-NDACC xCH4 smooth 100km 3hr					S5P-NDACC xCH4 smooth 100km bc 3hr					
site	#	rel. NDACC std	correlation	rel diff bias(%)	rel diff std(%)	#	rel. NDACC std	correlation	rel diff bias(%)	rel diff std(%)	lat
NY.ALESUND	4	2.7	0.89	3.57	1.07	4	2.7	0.97	4.54	1.02	78.9
THULE	54	1	0.75	3.48	0.75	54	0.9	0.78	4.49	0.73	76.5
KIRUNA	93	0.9	0	-1.13	1.5	93	0.8	0.01	-0.15	1.53	67.8
SODANKYLA	87	0.9	0.2	-0.7	1.12	87	0.9	0.24	0.33	1.12	67.4
ST.PETERSBURG	87	1.1	0.28	-0.63	0.95	87	1.2	0.39	0.31	0.86	59.9
BREMEN	30	1.4	0.66	0.74	0.69	30	1.4	0.71	1.32	0.66	53.1
GARMISCH	17	1.1	0.73	-0.39	0.54	17	1.1	0.76	0.48	0.52	47.5
ZUGSPITZE	33	0.9	0.75	-0.12	0.6	33	0.9	0.74	0.72	0.64	47.4
JUNGFRAUJOCH	25	0.8	0.61	-0.7	0.65	25	0.8	0.65	-0.04	0.65	46.6
TORONTO	67	3.7	0.27	-0.2	2.44	67	3.4	0.21	0.38	2.51	43.6
RIKUBETSU	12	2.4	0.24	0.81	1.99	12	2	0.17	1.6	2.11	43.5
BOULDER.CO	29	1.1	-0.07	2.23	1.02	29	1	0.12	2.05	0.96	40
ALTZOMONI	41	1	0.32	2.69	0.72	41	0.9	0.53	2.48	0.63	19.1
LAUDER	42	1.3	0.78	0.05	0.63	42	1.3	0.74	0.56	0.68	-45
ARRIVAL.HEIGHTS	11	0.7	0.68	1.66	0.99	11	0.8	0.59	2.45	1	-77.8
		1.4	0.47	0.76	**1.05**		1.3	0.51	1.44	**1.04**	

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.).

10.3.4 Dispersion

The 1σ spread of the relative difference (between the S5P and the TCCON methane column data) around the mean value is below the mission requirements (precision <1%) for both the bias corrected and standard products. The individual values for the different sites are indicated in **Figure 60**.

10.3.5 Dependence on influence quantities

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

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

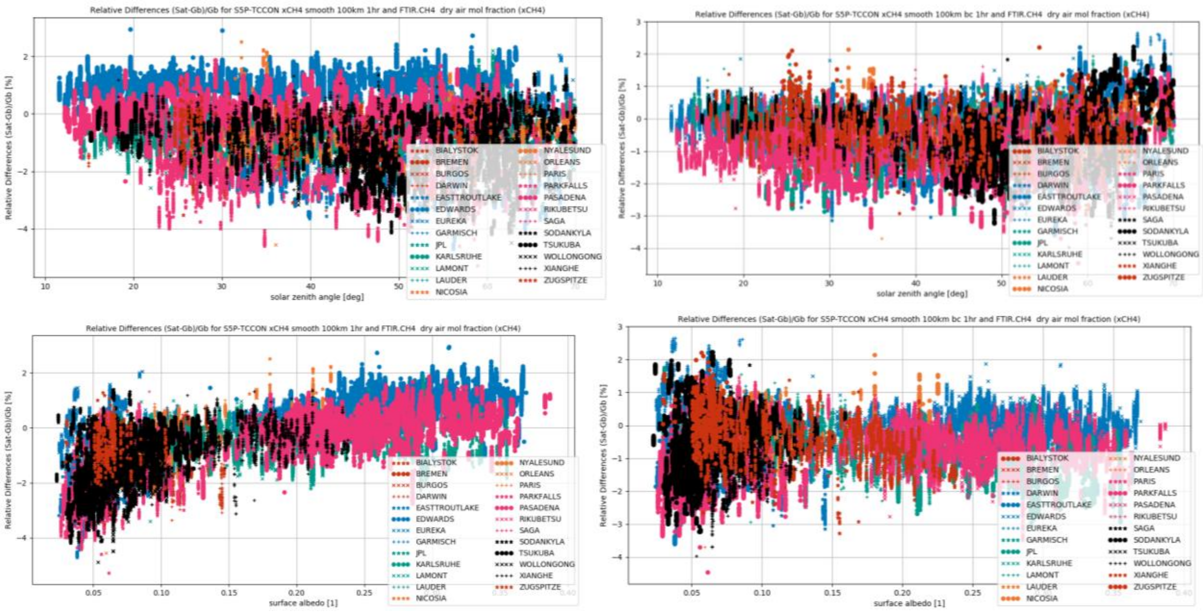


Figure 61: Dependence of the S5P-TCCON relative difference on solar zenith angle (top) and surface albedo (bottom). The left column shows the standard S5P product and the right column the bias corrected S5P product. The bias correction removes the surface albedo dependence of the standard S5P product.

10.3.6 Short term variability

For all the NDACC and TCCON stations, short scale temporal variations in the CH_4 column as captured by ground-based instruments are reproduced very similarly by S5P L2_CH4 OFFL. The individual Pearson correlation coefficients are on average 0.6, see **Figure 62**. At some sites the correlations are very low (e.g. Sodankylä, Burgos, 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.

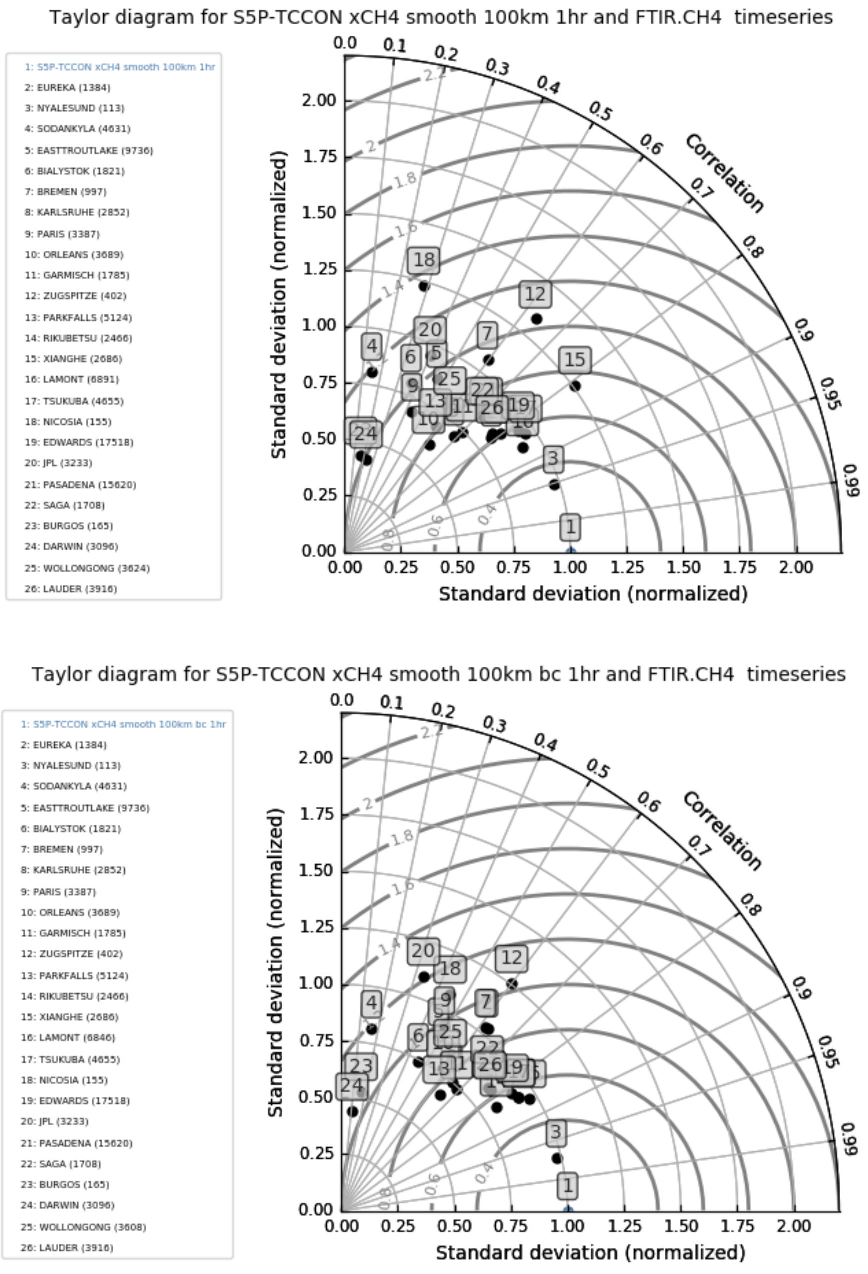


Figure 62: Taylor diagrams for differences between S5P L2_CH4 RPRO/OFFL and TCCON methane column data: standard (top) and bias corrected (bottom) S5P methane columns. At almost all sites the variability of the S5P column data is higher compared to the variability in the TCCON data.

10.3.7 Geographical patterns

Single TROPOMI overpasses show stripes of erroneous CH_4 values in the flight direction (see **Figure 63** left). For orbits before orbit 7644 (April 5 2019) not all pixels above inland water are filtered with the `qa_value` flag, see **Figure 63** (right, above Caspian Sea).

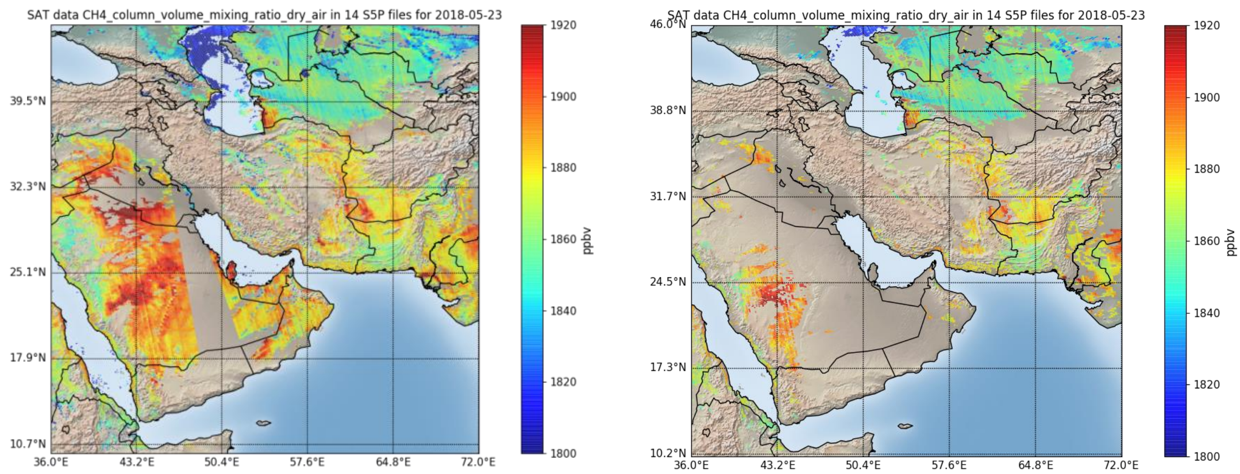


Figure 63: Maps showing XCH_4 concentrations above the Middle East measured on May 23 2018. The left panel shows all available pixels, the right panel shows only pixels with qa_value>0.5. The left panel shows the presence of stripes in the flight direction and the right panel shows the presence of filtered pixels above inland water (Caspian Sea).

10.3.8 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.

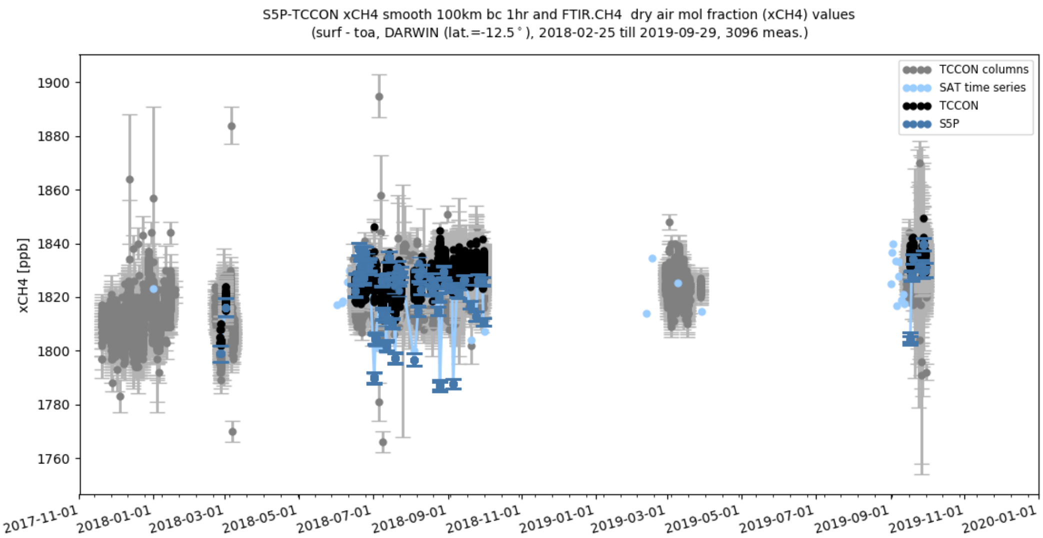


Figure 64: S5P L2_CH4 XCH_4 time series over Darwin where low values of XCH_4 are observed for several days.

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**. The NRTI and OFFL processors are currently based on the same algorithm and are producing very similar data products; therefore, only validation of the L2_CLOUD OFFL product is reported hereafter. Subsection 11.4 demonstrates evidence that NRTI and OFFL data do not differ significantly and that their respective validations yield similar conclusions

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 17 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.

Station	Location	Network	Latitude (N)	Longitude (E)
Ny-Ålesund	Svalbard	CLOUDNET	78.932	11.921
Summit	Greenland	CLOUDNET	72.60	-38.42
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
Iquique	Chile	CLOUDNET	-20.54	-70.18
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: 15 CLOUDNET sites and 2 ARM (Atmospheric Radiation Measurement) sites. Data is collected from EVDC.

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.

11.2.2 Satellites

MODIS and NPP VIIRS

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 VIIRS NASA non-operational product[1]. The comparison with MODIS allows only for daily means validation whereas the comparison against VIIRS offers a pixel-by-pixel validation of the product.

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).

[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 have been released in Feb. 2019 and 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.

11.2.3 Alternative S5P cloud algorithms

S5P FRESCO

The support product S5P TROPOMI FRESCO cloud height is also compared to CLOUDNET observations, and directly with CLOUD CRB at the CLOUDNET sites. This helps to judge if discrepancies between S5P CLOUD CRB and CLOUDNET are specific to the adopted cloud retrieval algorithm or are of more general nature. The S5P FRESCO support product is not publicly disseminated separately, but is used as input for e.g., the S5P NO2 retrieval. Earlier versions of the algorithm are described in e.g., [Koelemeijer 2001]. Like CLOUD CRB, FRESCO-S models a cloud as a Lambertian reflector. Information on cloud pressure and effective cloud fraction is derived from the reflectance in and around the O_2 A band. As opposed to CLOUD CRB, where cloud albedo is retrieved, in FRESCO-S, the cloud albedo is assumed to be 0.8 or the TOA reflectance at 758 nm if the reflectance is larger than 0.8. We note that at small cloud fractions, the surface albedo is adapted to prevent negative cloud fractions.

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*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^{o}), 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) and at the *3rd S5PVT 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, while up-to-date validation results and consolidated validation reports are available through the MPC VDAF Portal at http://mpc-vdaf.tropomi.eu.

Current conclusions are based on the limited amount of reference measurements available at the time of this analysis, resulting in a limited amount of comparison pairs. The validation vs. CLOUDNET ground-based data uses S5P L2_CLOUD RPRO+OFFL 01.01.07 data. This covers the time period from 2018-04-30 till 2019-11-16. CLOUDNET data from 17 sites were considered in this analysis.

11.3.3 Radiometric cloud fraction (L2_CLOUD CAL & L2_CLOUD CRB)

11.3.3.1 Bias

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 (CRB-FRESCO)/[0.5*(CRB+FRESCO)].

Figure 65 presents boxplots of the difference (top) and normed relative difference (bottom) between rescaled cloud fractions of CLOUD CRB and FRESCO. Overall, mean and median difference are negative (CRB lower than FRESCO) but within the 20% bias requirement. The strong deviations at Summit and Ny-Ålesund are at least partly due to pixels with snow-ice cover.

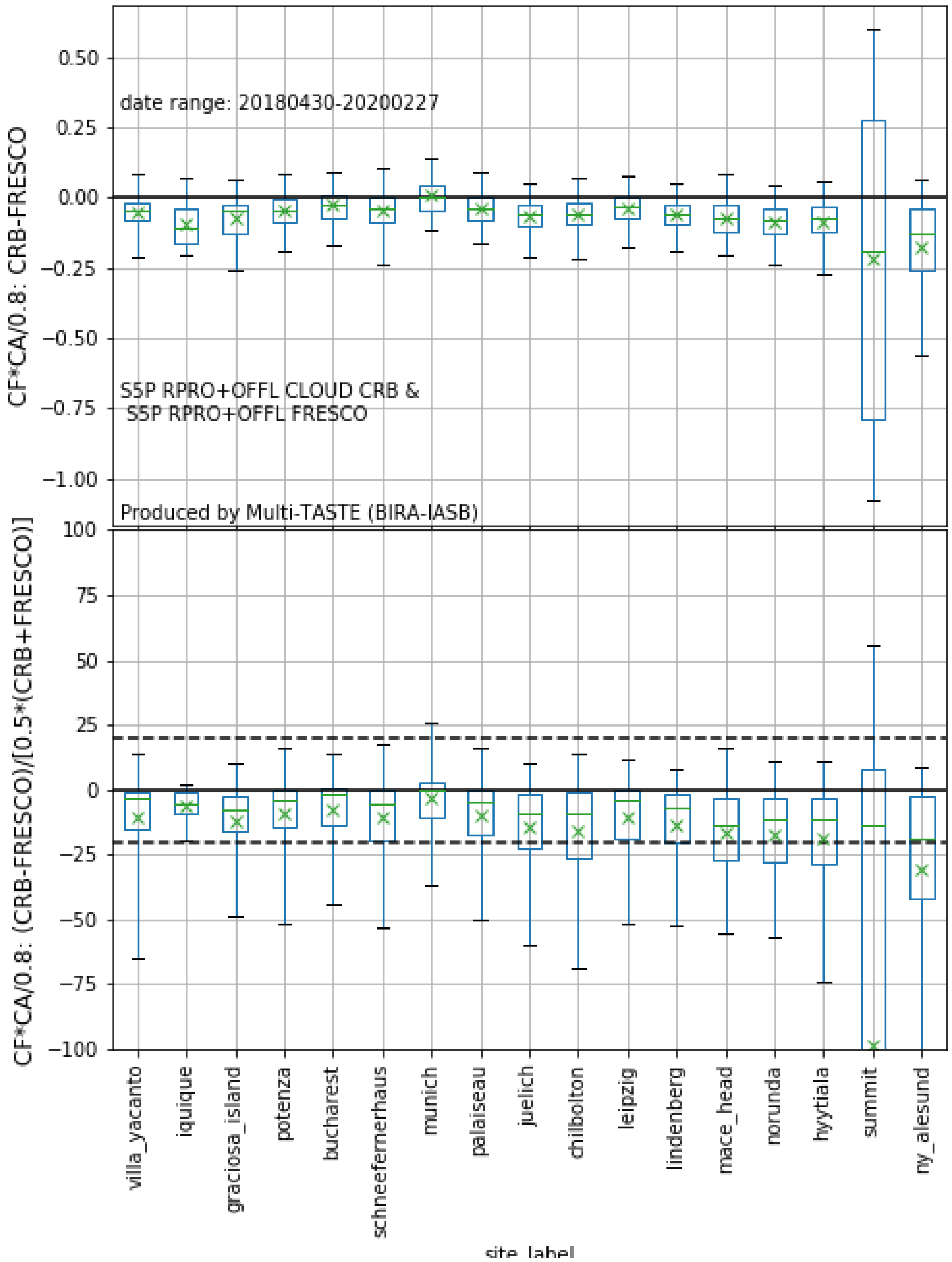


Figure 65. Top. Box plots of S5p CLOUD CRB cloud fraction minus S5p FRESCO cloud fraction, after both have been rescaled to a cloud albedo of 0.8. Bottom: same but for the normed relative difference.

11.3.3.2 Dispersion

Figure 66 presents two dispersion measures of the CRB minus FRESCO rescaled cloud fractions: standard deviation (left) and the more robust ½ IP68. While the standard devation is close to 0.1 at most sites, the ½ IP68 is lower and only slightly exceeding the dispersion requirement of 0.05.

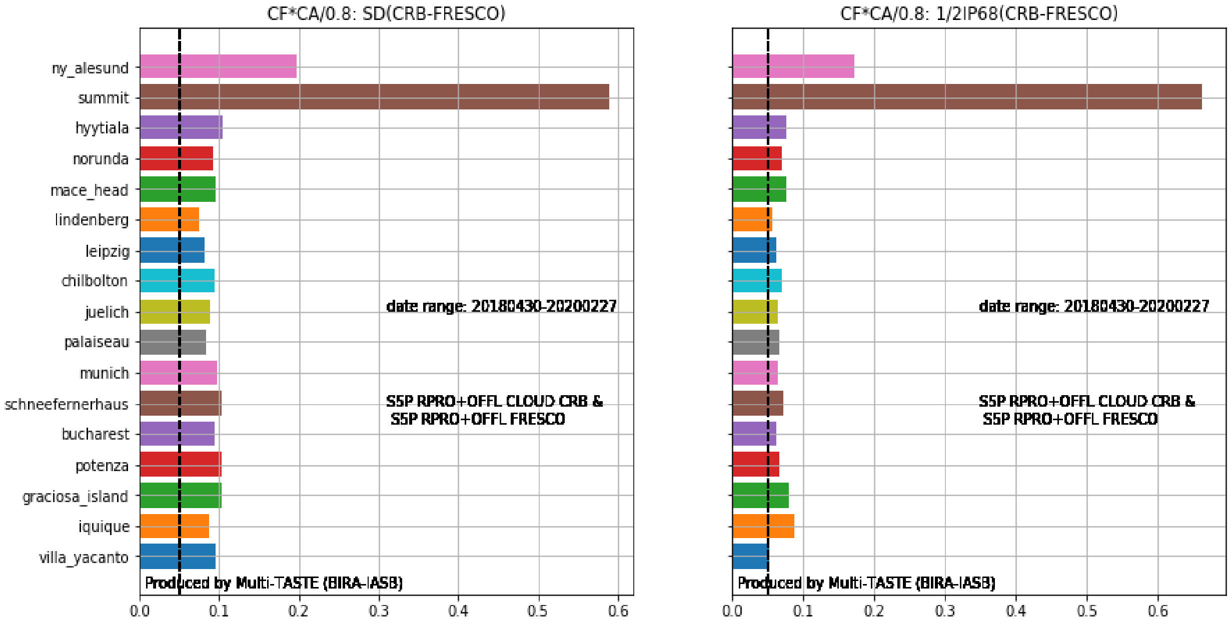


Figure 66. Left. Standard deviation S5p CLOUD CRB cloud fraction minus S5p FRESCO cloud fraction, after both have been rescaled to a cloud albedo of 0.8. Right: same but for the ½ IP68.

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 new L1b version 2 data which contain 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.x.x (not yet operational). 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 **Figure 67**.

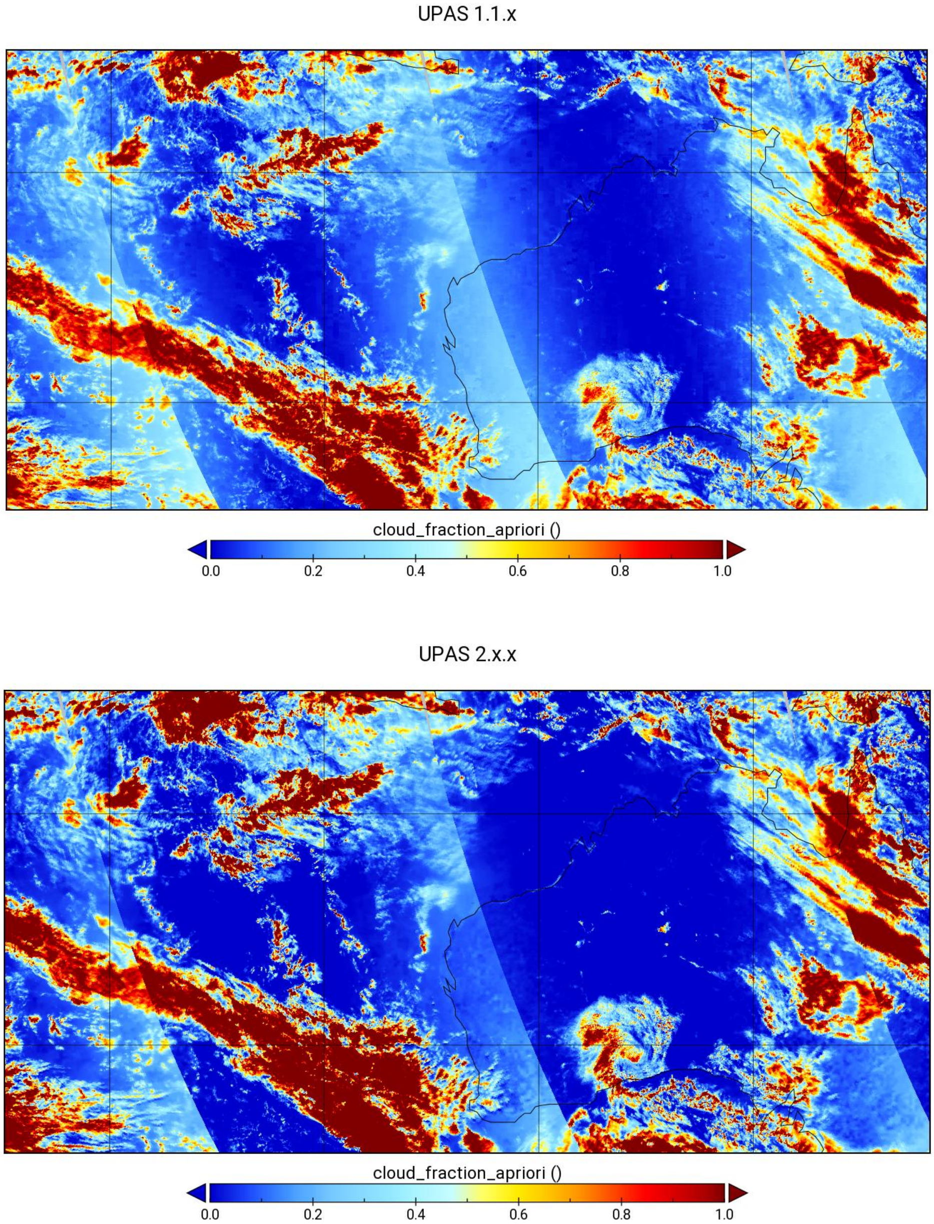


Figure 67. Top. S5p CLOUD OCRA a priori cloud fraction of the operational product L2_CLOUD RPRO 1.1.5. Bottom. The same, but for the (not yet operational) UPAS 2.x.x. Orbits 03614, 03615, 03616 and 03617 on 2018-06-25.

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 68, Figure 69). The monthly mean S5P CAL CTH generally follows the trend of the CLOUDNET cloud top height (Figure 68). This is corroborated by the high Pearson correlation coefficient of 0.8 (Figure 69). In absolute scale terms, the overestimation is higher for high clouds (Figure 69). On the other hand, Figure 69 makes clear that the 20% upper limit requirement on the bias becomes very strict for low clouds. Note that with the second validation approach, there are more near-surface CLOUD CAL CTH values.

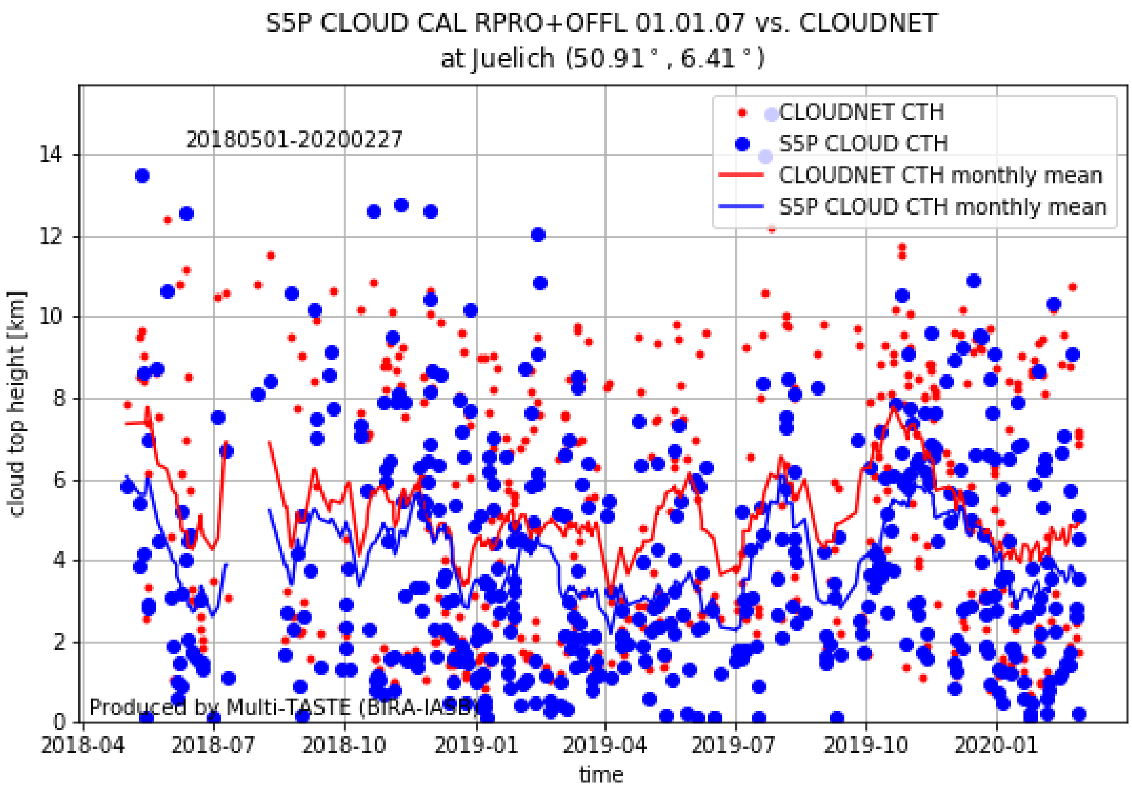


Figure 68: Time series of S5P CLOUD CAL (RPRO and OFFL, processor version 01.01.07) CTH vs. CLOUDNET CTH at Jülich. The monthly mean of both is also provided. Sensing time range is indicated on the figure.

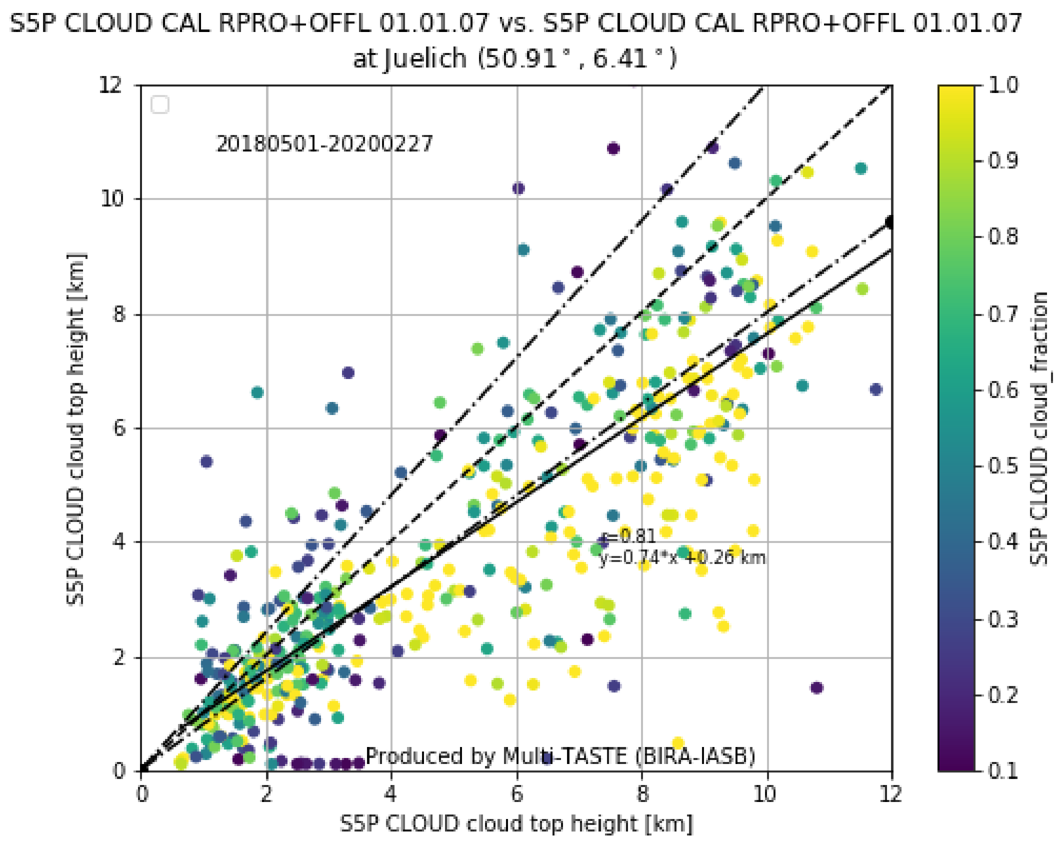


Figure 69: Correlation plot of S5P L2_CLOUD CAL (RPRO+OFFL, processor version 01.01.07) CTH vs. CLOUDNET CTH at Jülich. The colour indicates the S5P L2_CLOUD cloud fraction. Dashed line is the 1:1 line and dash-dotted line the 20% bias requirement.

At most of the sites, the S5P L2_CLOUD CAL top height is lower than the CLOUDNET top height, as depicted in **Figure 70**.

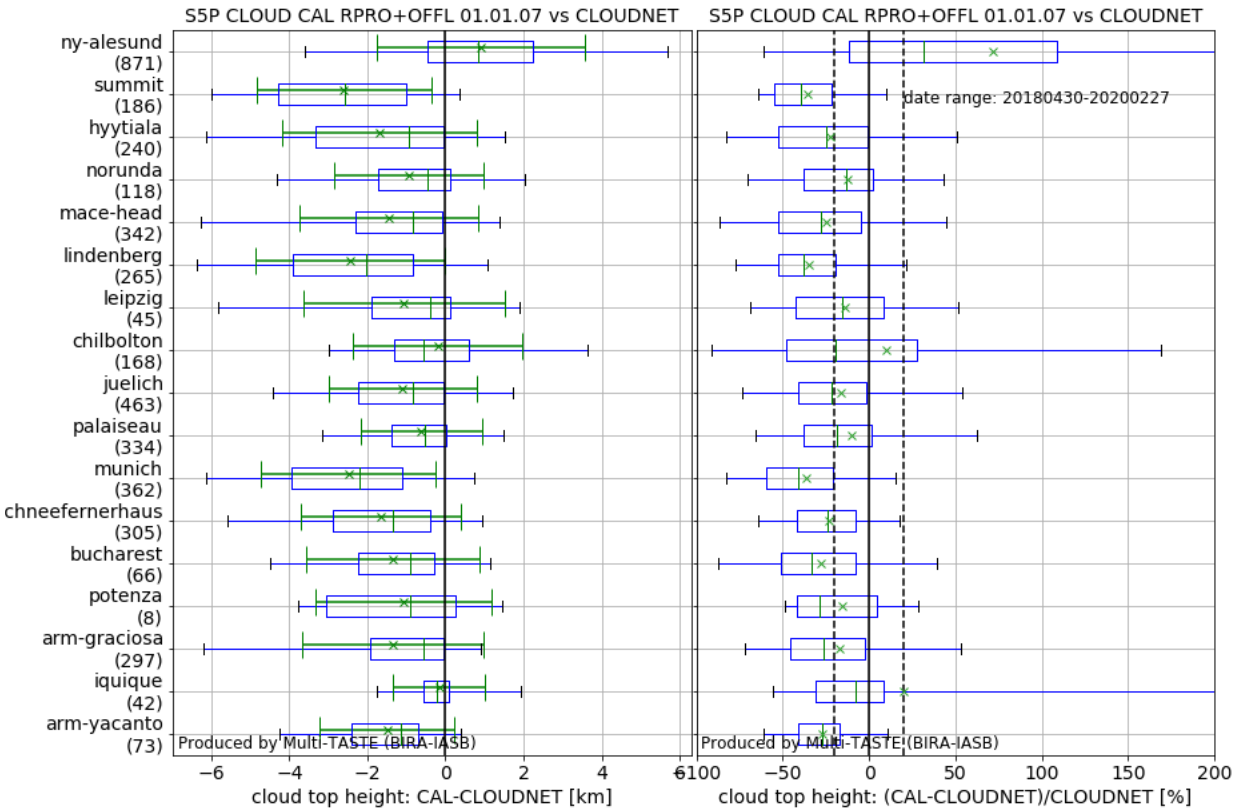


Figure 70: Left panel: Boxplots of S5P L2_CLOUD CAL CTH minus CLOUDNET CTH, per site. Boxplot conventions: box bounds are at first and third quartile. Green line is median. Whiskers are at 5 and 95 percentiles. Slightly offset green cross is mean, with green error bar ±the standard deviation. Right panel: 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 cover (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. Possibly snow/ice cover also plays a role here. This should be further investigated.
- Regarding Chilbolton, there is a large discrepancy between the mean relative difference (+15%) and the median relative difference (-20%). The large mean relative difference is caused by the presence of a limited amount of comparison pairs where CLOUDNET reports a low cloud top height (< 1 km) while S5P L2_CLOUD CAL reports a much higher cloud top

height (several km). Differently from the Ny-Ålesund case, there is no clear link with retrieved CLOUD CF. This should be further investigated.

- The cause of the exceptional large deviations for Munich and Lindenberg has to be further investigated (e.g., role of orography, cloud type, land cover…).

CLOUD CRB cloud height is generally below the CLOUDNET cloud height. A typical case is provided for the CLOUDNET site at Jülich (Figure 71, Figure 72). The monthly mean S5P CRB CH generally follows the trend of the CLOUDNET cloud mean height (Figure 71). This is corroborated by the high Pearson correlation coefficient of 0.81 (Figure 72). Figure 72 indicates that the 20% upper limit requirement on the bias is more easily met for high-lying clouds.

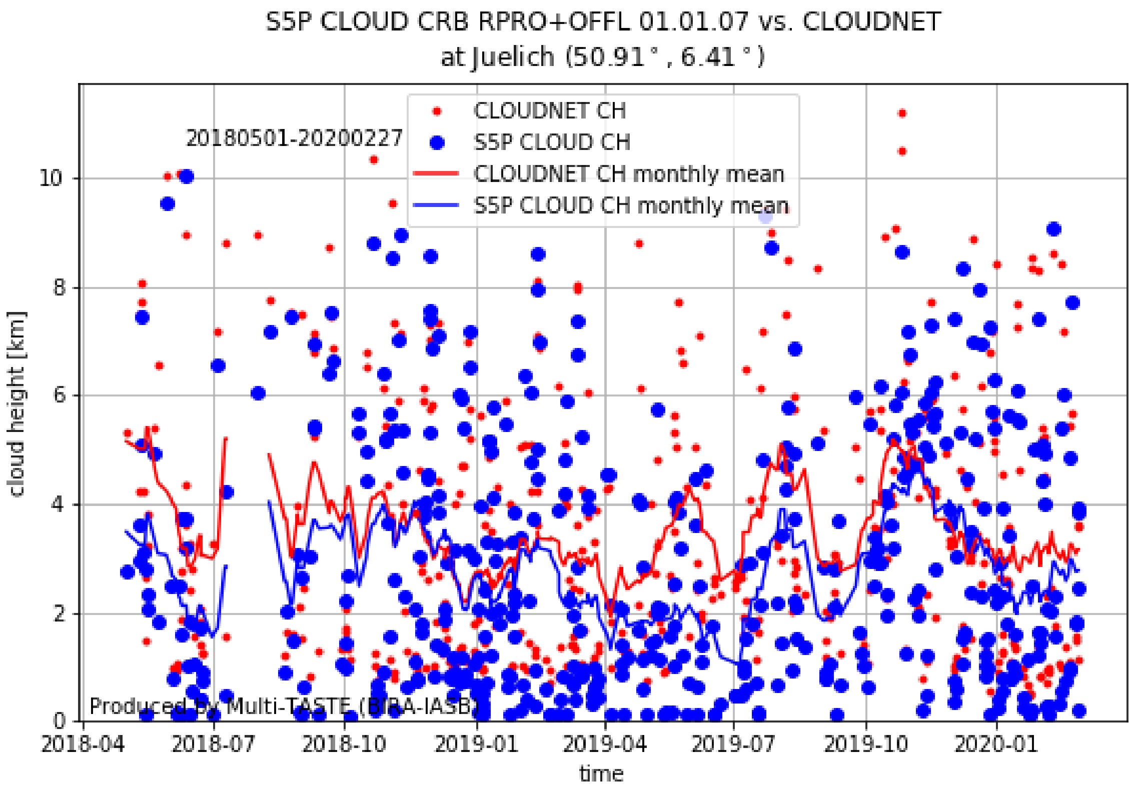


Figure 71: Time series of S5P L2_CLOUD CRB (RPRO and OFFL, processor version 01.01.07) CH vs. CLOUDNET CH at Jülich. The monthly mean of both is also provided. Sensing time range is indicated on the figure.

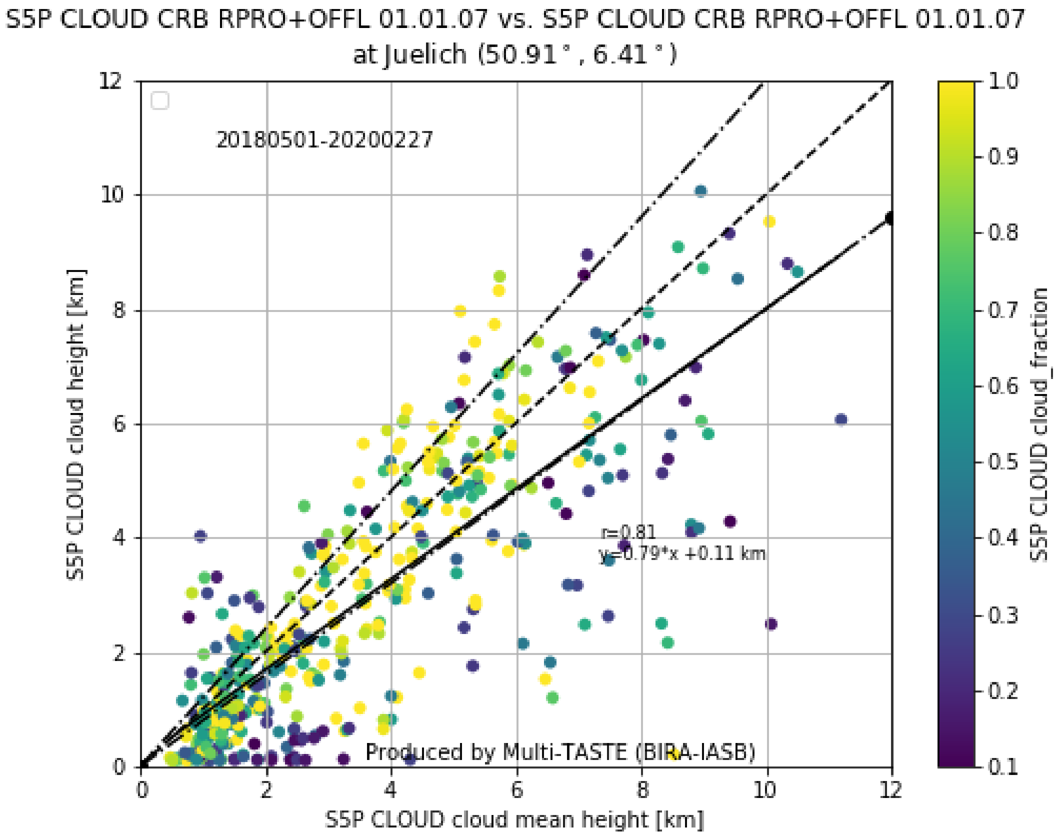


Figure 72: Correlation plot of S5P L2_CLOUD CRB (RPRO and OFFL, processor version 01.01.07) CH vs. CLOUDNET cloud mean height at Jülich. The colour indicates the S5P CLOUD cloud fraction. Dashed line is the 1:1 line and dash-dotted line the 20% bias requirement. Sensing time range is indicated on the figure.

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). Other exceptions (e.g., Munich, Lindenberg) deserve further investigation. 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 73) while this is not the case for FRESCO vs CLOUDNET. We return to this in Section 0.

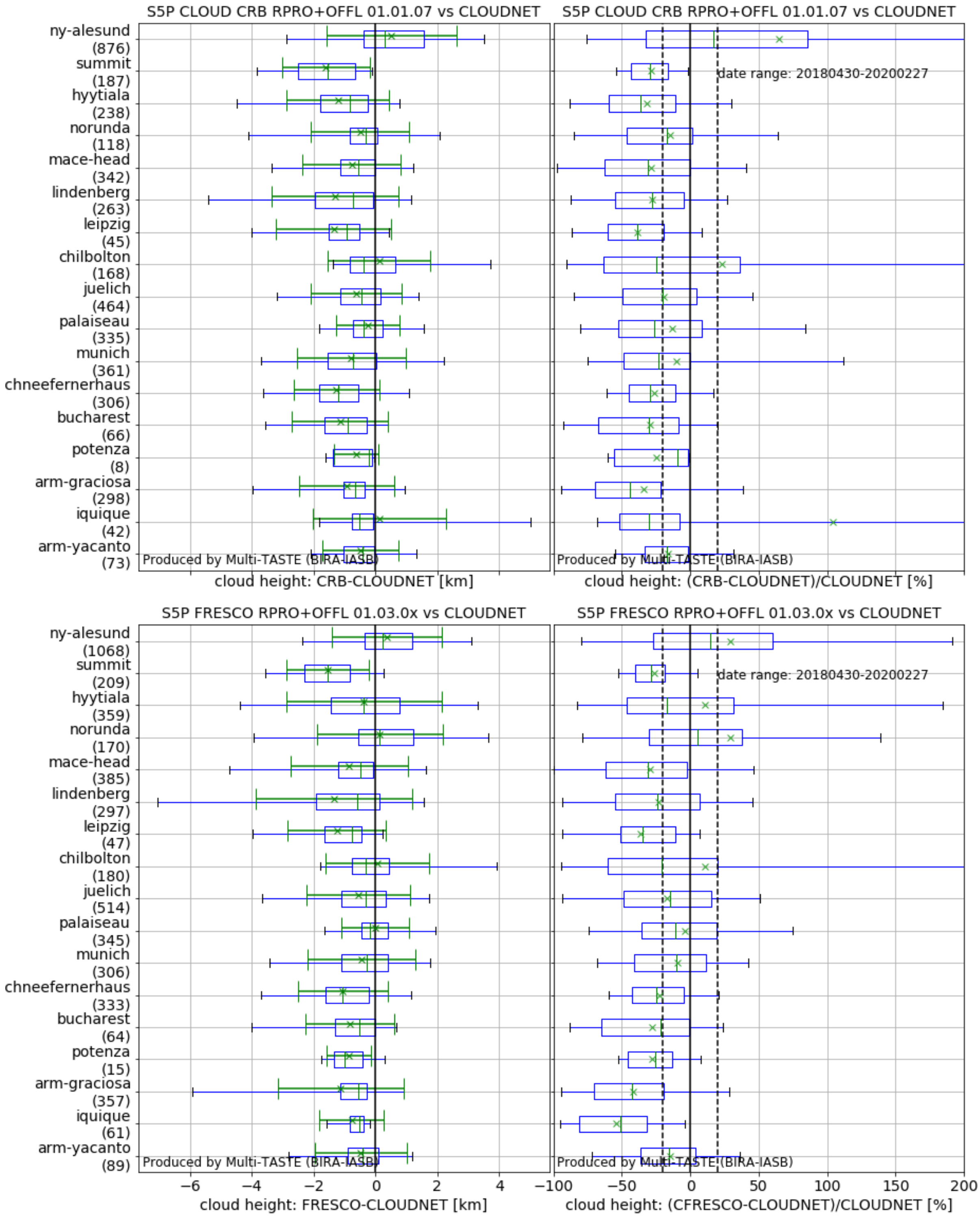


Figure 73: Upper panel: Boxplots of S5P L2_CLOUD CRB CH minus CLOUDNET CH (upper left) and of the relative difference (upper right), per site. Lower panel: The same but now for the S5P Support product FRESCO RPRO 01.03.02 + OFFL 01.03.00-02. The same conventions as for **Figure 70** apply. Sensing time range is indicated on the figure.

Comparison with alternative S5p cloud height retrievals

S5p CLOUD CRB CH was compared with S5p FRESCO 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 FRESCO) occur at the sites Iquique, Ny-Ålesund and Summit.

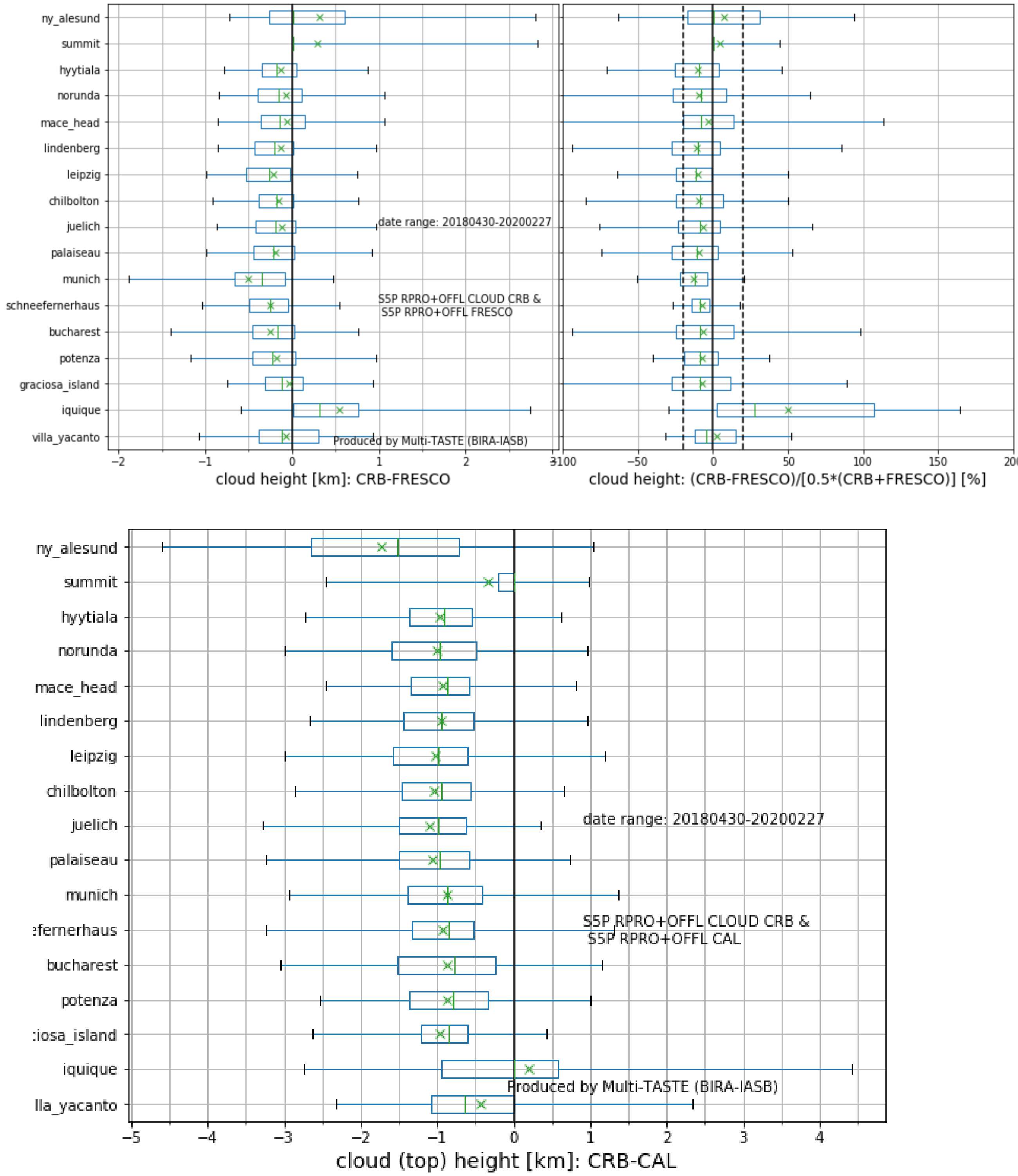


Figure 74: Top. Boxplots of S5p CLOUD CRB CH minus S5p FRESCO CH (left) and of the normed difference (right). Bottom. Boxplots of S5p CLOUD CRB CH minus S5p CLOUD CAL CTH.

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.

Comparison with MODIS and NPP VIIRS

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 (**Figure 75**).

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

Possible bias in the cloud fraction is difficult to be identified because of the comparison of radiometric cloud fraction from TROPOMI against geometrical cloud fraction from MODIS/VIIRS.

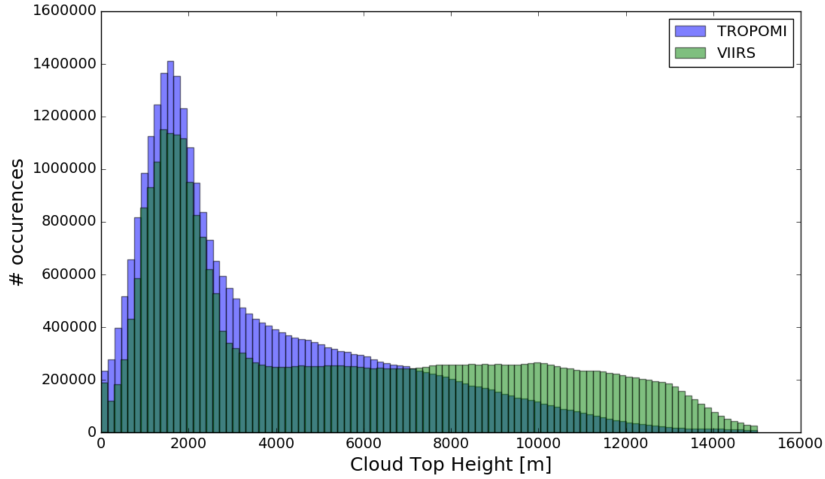


Figure 75: Histograms of cloud top height of S5P L2_CLOUD (internal prototype comparable to operational processing version 01.01.05) cloud top height and VIIRS cloud top height.

11.3.4.2 Dispersion

Comparison with CLOUDNET cloud top height and cloud height

From **Figure 70** 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 73**). 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.

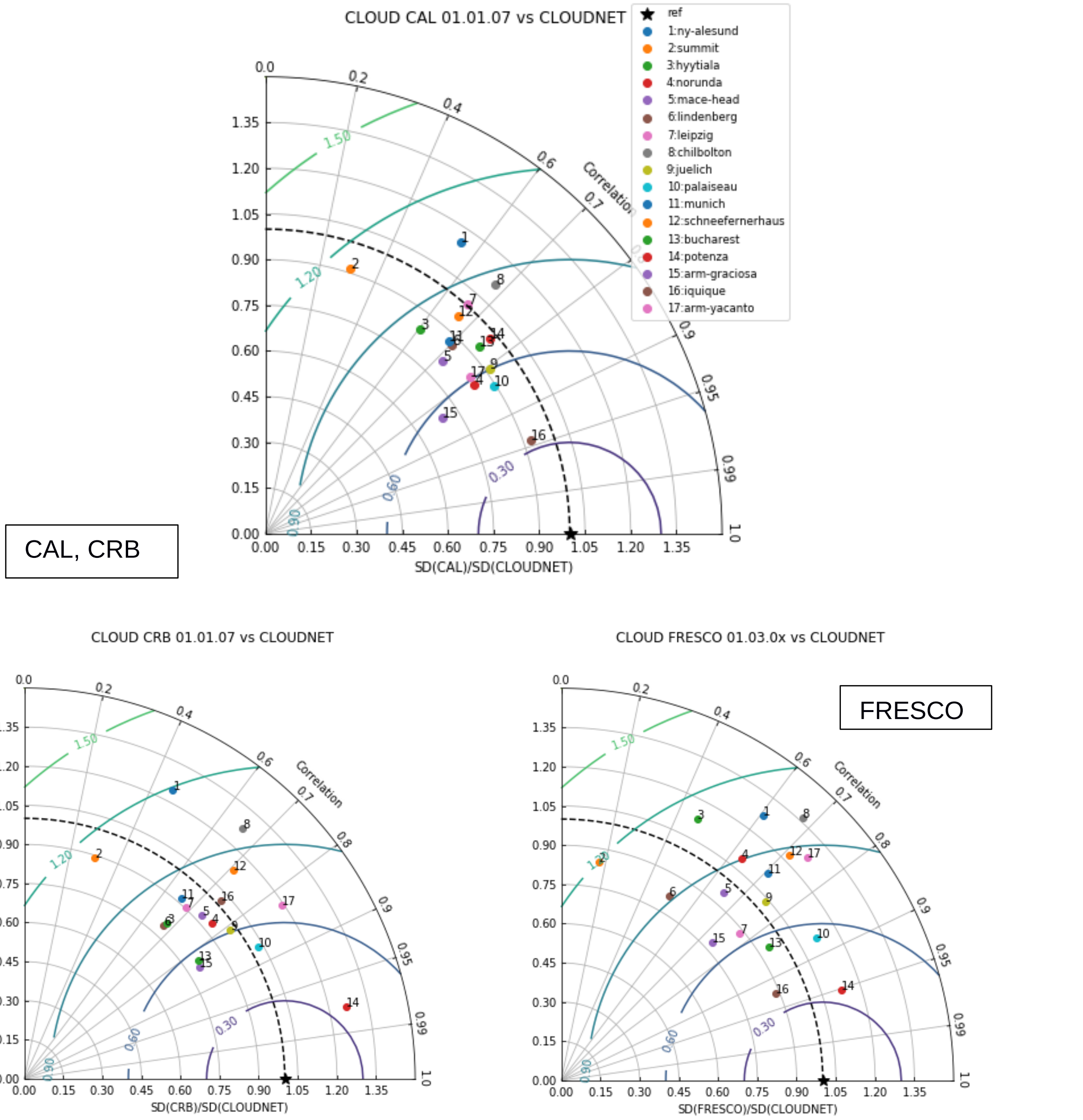


Figure 76: Taylor diagrams of CLOUD CAL vs CLOUDNET CTH, CLOUD CRB vs CLOUDNET CH and FRESCO CH vs CLOUDNET CH. The radius represents SD(S5p)/SD(CLOUDNET), the angle represents correlation and the circular contour lines represents the relative dispersion SD(S5p-CLOUDNET)/SD(CLOUDNET).

Figure 76 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, Chilbolton, and specifically for FRESCO, Hyytiala. 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 77**, left). Exceptions are Iquique, Summit, Munich and Ny-Alesund where the dispersion is considerably higher. The more robust dispersion measure ½ 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 74**, 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.

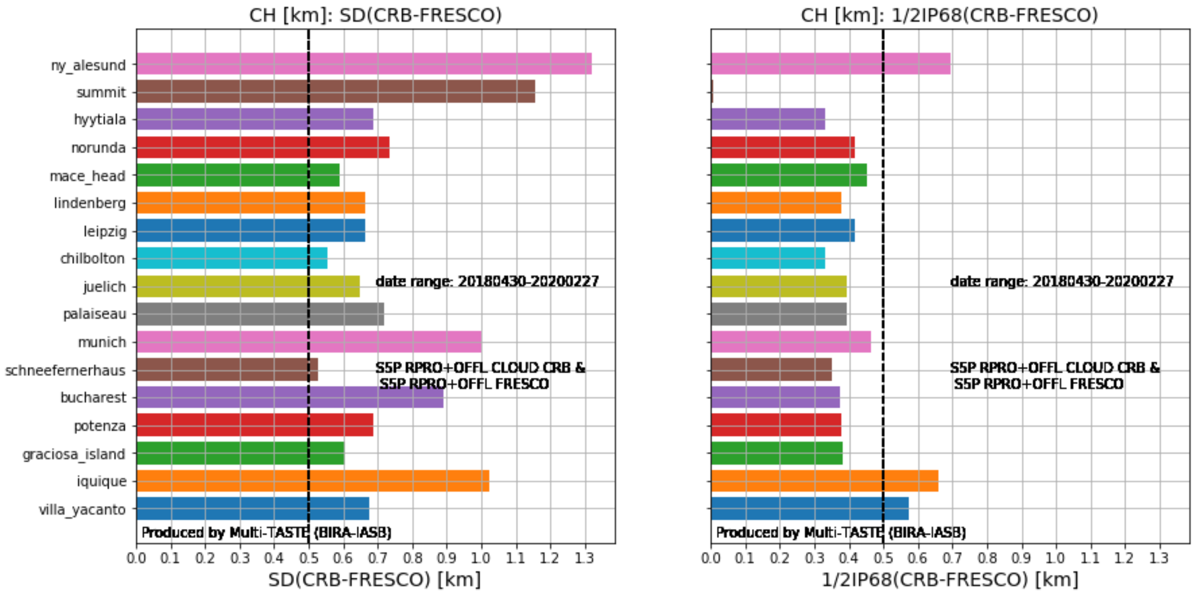


Figure 77: Standard deviation (left) and ½ IP68 (right) of S5p CLOUD CRB CH minus S5p FRESCO CH.

Comparison with NPP VIIRS

S5P L2_CLOUD CTH shows good correlation with VIIRS CTH: Pearson coefficient R = 0.74 for continental clouds and 0.86 for marine clouds (see **Figure 78**).

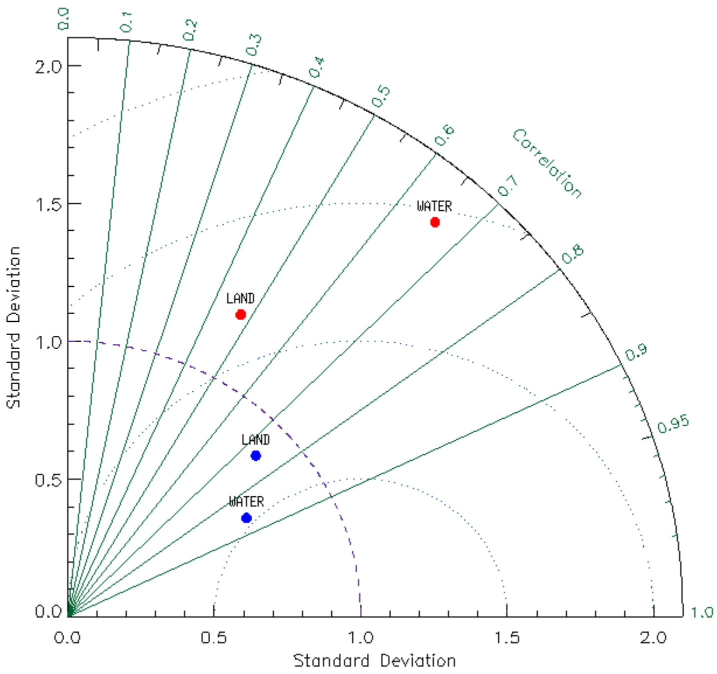


Figure 78: Taylor diagram between CTH (blue) and COT (red) of S5p CLOUD and those of VIIRS.

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 74**, top), CLOUD CRB CH – CLOUD CAL CTH (**Figure 74**, bottom), and CLOUD CRB CH – CLOUDNET (**Figure 73**). 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 79** 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 79** (top) are closer to the 1:1 line than the FRESCO data points in the 0-2km x 0-2km box of **Figure 79** (bottom).

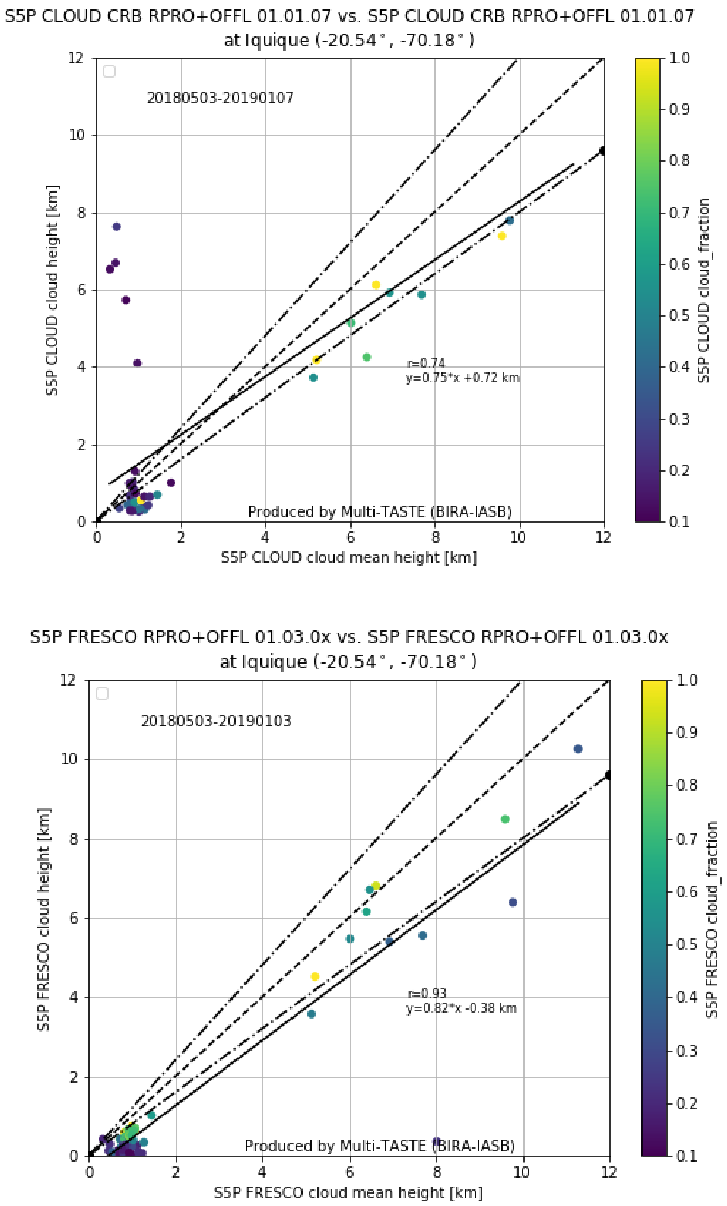


Figure 79. Top. S5p CLOUD CRB CH vs CLOUDNET CH correlation plot. Bottom. S5p FRESCO CH vs CLOUDNET CH correlation plot.

Comparison with alternative S5p cloud height retrievals

Figure 80 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.

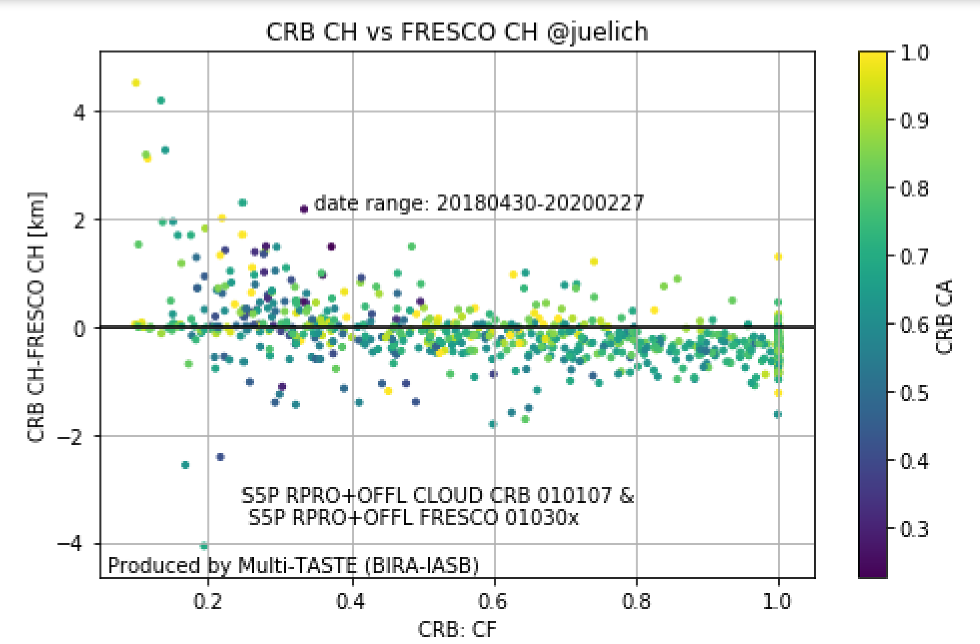


Figure 80: Cloud height difference between S5p CLOUD CRB and S5p FRESCO, in function of CLOUD CRB cloud fraction. Color scale is CLOUD CRB cloud albedo.

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. The effect seems to be strongest in the latitude bands [40-60]N and [30-40]S. This issue is reduced in UPAS 2.x.x (not yet operational). An example for CLOUD CAL cloud top height is shown in **Figure 81**.

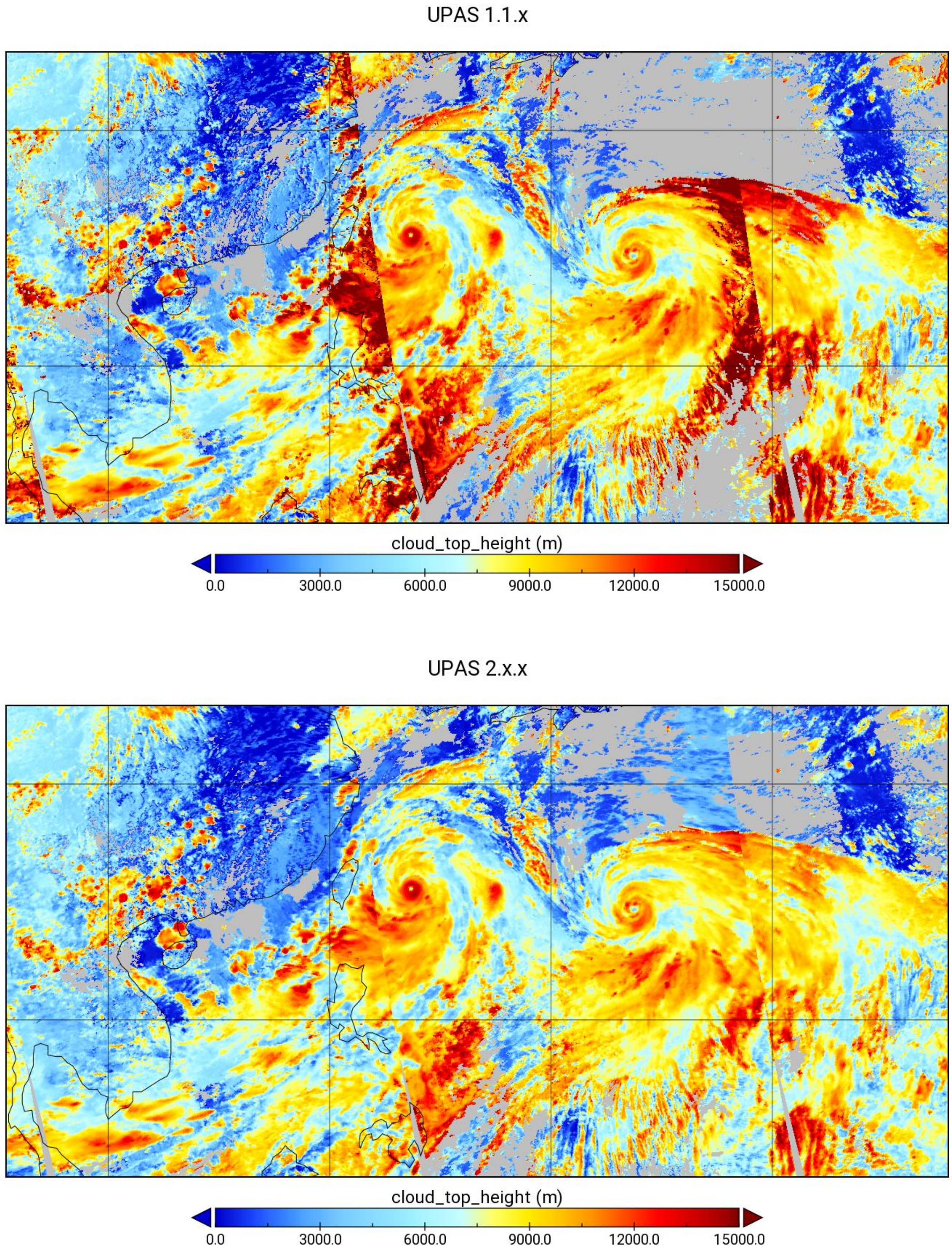


Figure 81. Top. S5p CLOUD CAL cloud top height of the operational product L2_CLOUD OFFL 1.1.7 (parts of orbits 09416, 09417, 09418 on 2019-08-08). Bottom. The same, but for the (not yet operational) L2_CLOUD 2.x.x. Note that enhancements near the east swath edge are reduced in L2_CLOUD 2.x.x.

11.3.4.6 A priori bug and its fix

Prior to version 01.01.06, a bug, causing cloud (top) height values being close to its a priori value of 3.8 km for a number of pixels (see Figure 82), impacted the CLOUD product quality. In processor version 01.01.06, this bug was corrected.

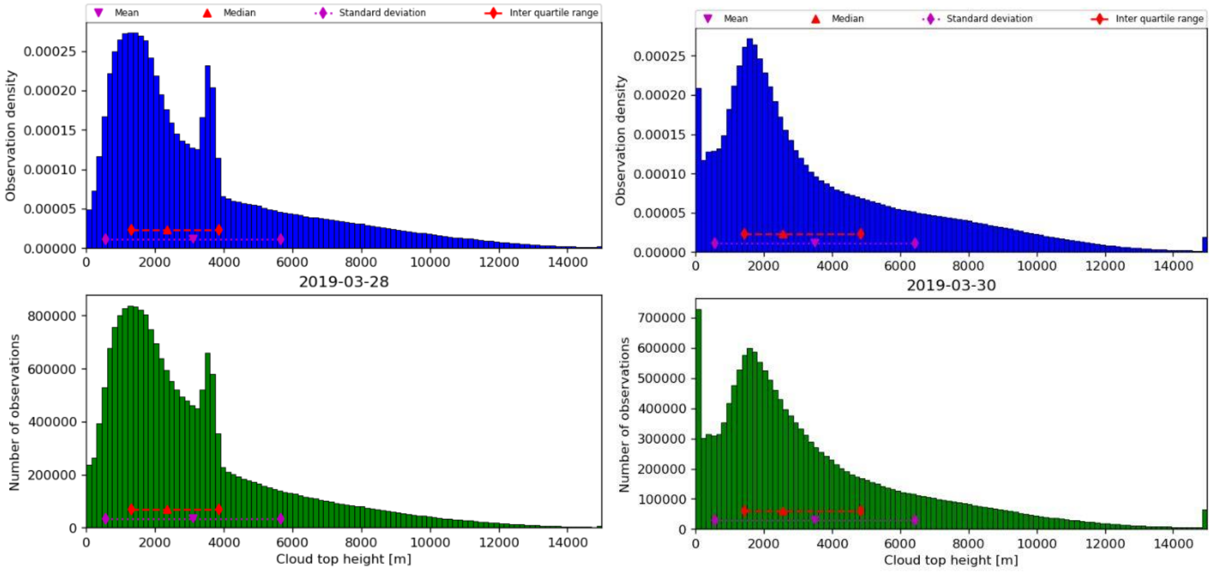


Figure 82: Histograms of cloud top height of S5P CLOUD. Left. Version 01.01.05, on the day before the version switch to 01.01.06. An artificial peak at 3.8 km, the a priori value, is visible. Right. Version 01.01.06. The peak at 3.8 km has disappeared.

11.3.4.7 Switch to smaller ground pixel resolution

The effect of the switch to a smaller ground pixel size on 6 August 2019 on the retrieved L2_CLOUD cloud parameters was investigated by comparing histograms (**Figure 84**) and global maps (**Figure 85**) on 5 and 6 August. No obvious effect on the retrieved values is noticed. Note that, as expected, the count of the histograms has increased.

Comparison with CLOUDNET cloud top height and cloud height

Comparisons with CLOUDNET at Jülich depicted in **Figure 83** are not noticeably affected.

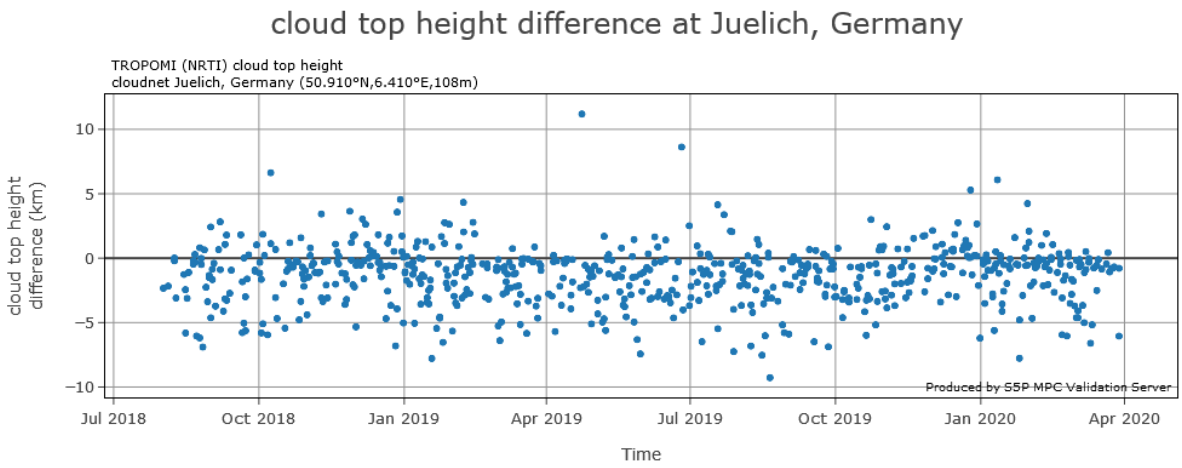


Figure 83: S5P NRTI CLOUD 01.01.01-01.01.07 CAL CTH vs CLOUDNET CTH at Jülich, as taken from the MPC Automated Validation Server.

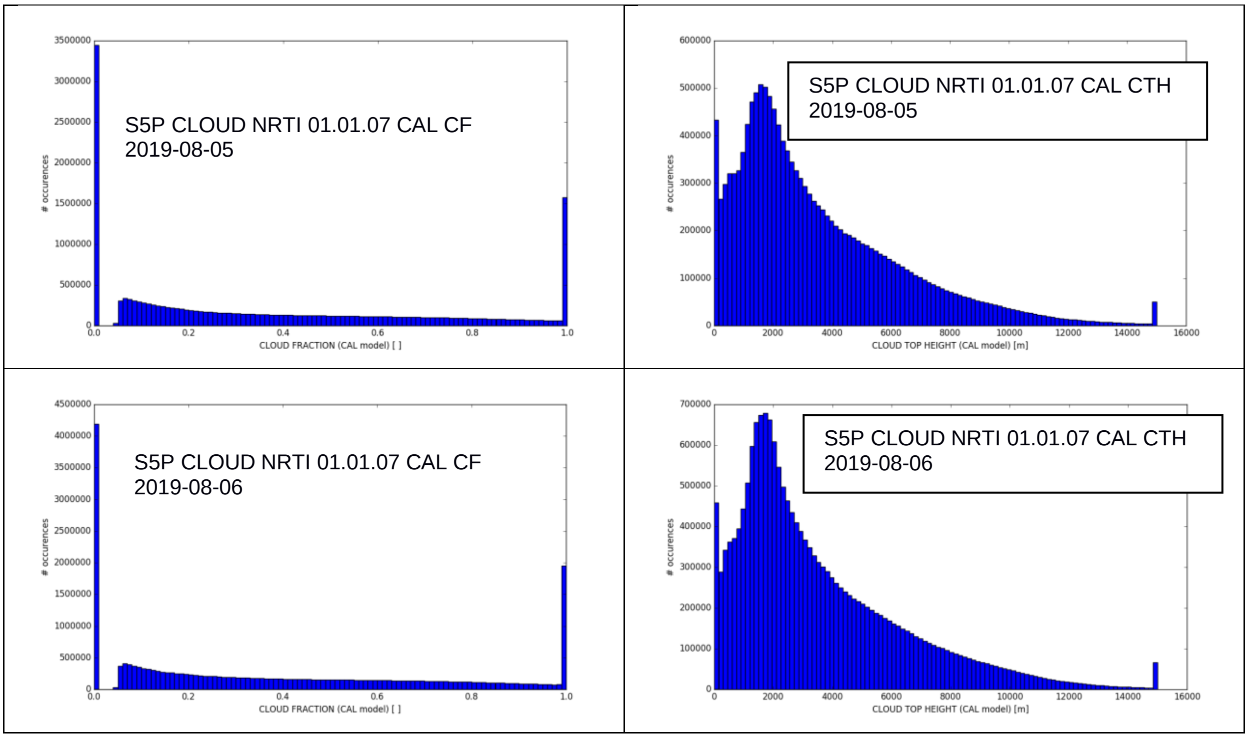
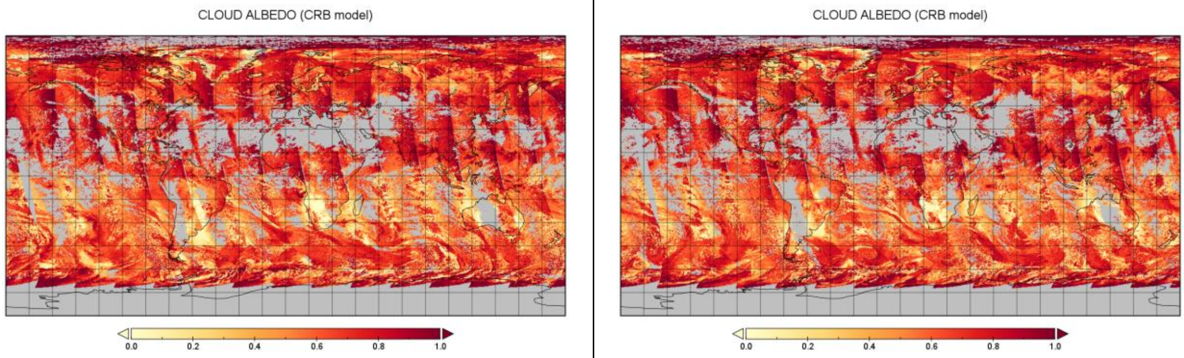


Figure 84: Histograms of S5P CLOUD NRTI 01.01.07 CAL cloud fraction (left) and CAL cloud top height (right), at 2019-08-05, before the pixel size switch (top) and at 2019-08-06, after the pixel size switch (bottom). The histograms shapes are very similar, but note that the count has increased. The same conclusion can be reached for the other CLOUD retrieved parameters.



S5P CLOUD NRTI 01.01.07 CRB CA 2019-08-05 | S5P CLOUD NRTI 01.01.07 CRB CA 2019-08-06

Figure 85: Global map of CLOUD NRTI 01.01.07 CRB cloud albedo at 2019-08-05 (before the pixel size switch) and at 2019-08-06 (after the switch).

11.3.5 Cloud optical thickness (L2_CLOUD CAL)

11.3.5.1 Bias

Comparison with NPP VIIRS

A positive bias (+7.9) in the cloud optical thickness (COT) with respect to VIIRS has been found.

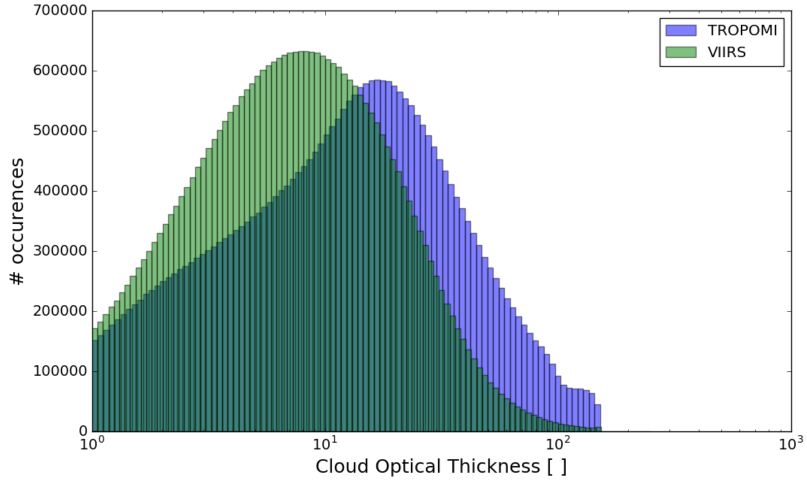


Figure 86: Histograms of cloud optical thickness of S5P L2_CLOUD (internal prototype comparable to operational processing version 01.01.05) and VIIRS cloud optical thickness.

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 -60ºS and +60ºS).

11.3.5.2 Dispersion

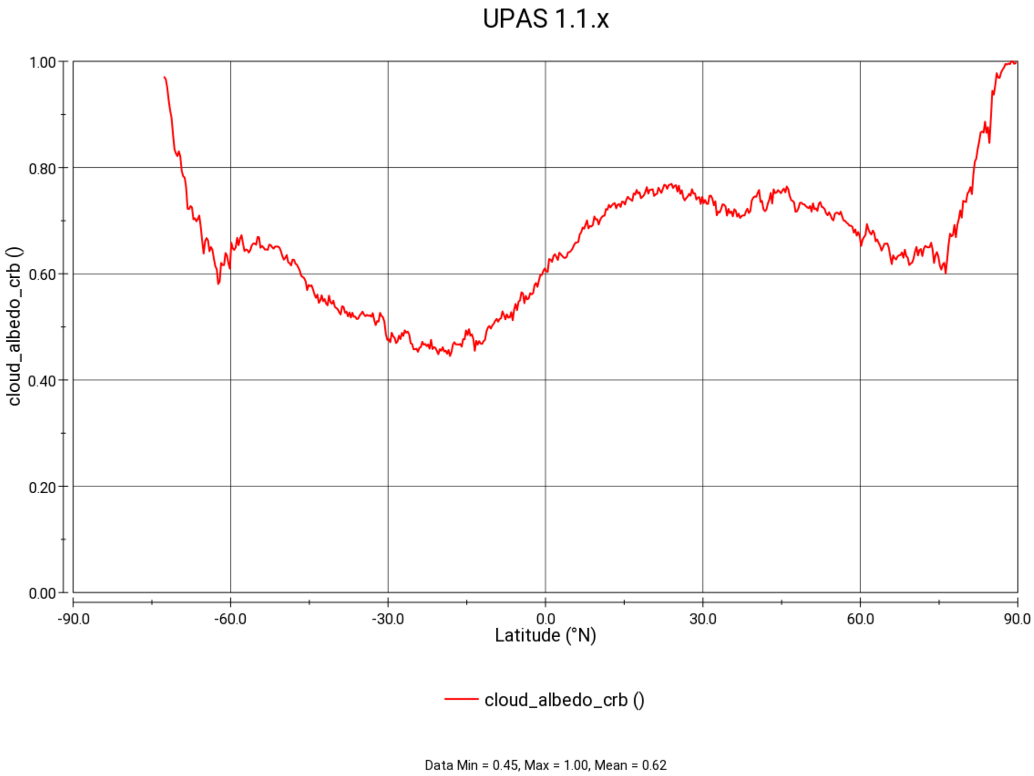
Comparison with NPP VIIRS

For COT, a weak correlation with NPP VIIRS is seen. Pearson R = 0.48 for continental clouds and 0.66 for marine clouds (see **Figure 78**).

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. This issue is reduced in UPAS 2.x.x (see **Figure 87**).



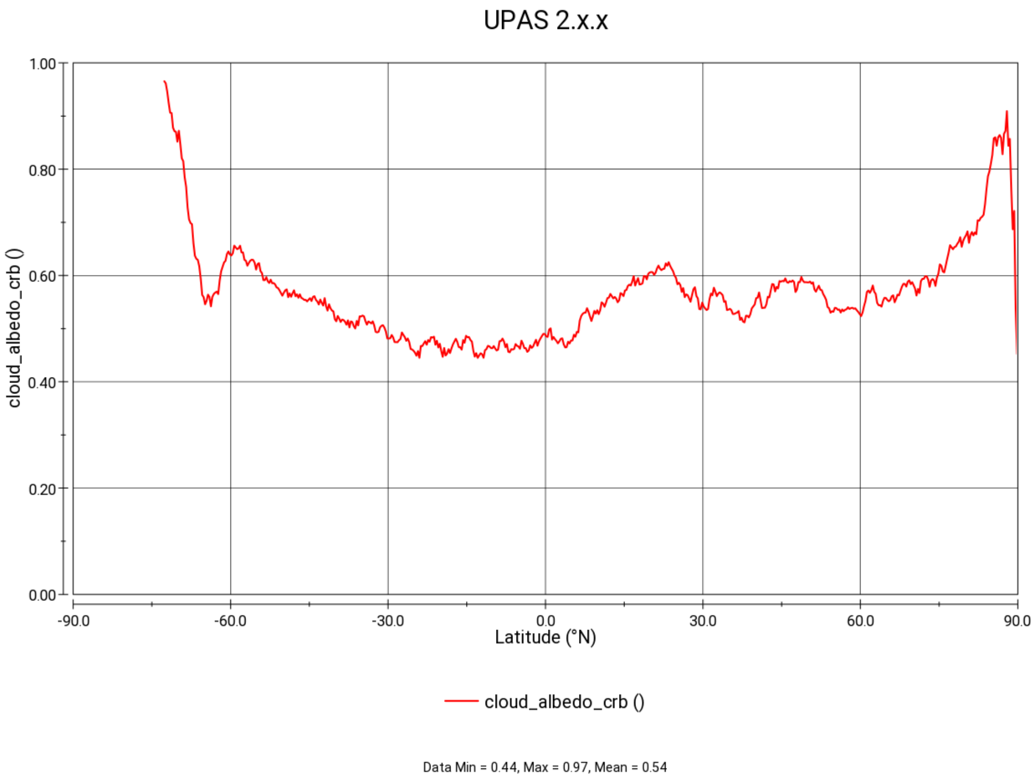


Figure 87: Top. S5p CLOUD CRB cloud albedo of the operational product L2_CLOUD 1.1.7, based on all orbits of 2019-08-08. Bottom. The same, but for the (not yet operational) L2_CLOUD 2.x.x. Note the reduced North-South gradient for L2_CLOUD 2.x.x.

11.4 Equivalence of L2_CLOUD NRTI and OFFL products

This section shows evidence that the L2_CLOUD NRTI and OFFL products do not differ significantly and that their respective validations yield similar conclusions.

CLOUD NRTI and OFFL use currently the same algorithm and therefore their difference is expected to be negligible. **Figure 88** shows illustrative demonstration for cloud top height, but similar conclusions apply to cloud fraction and cloud optical thickness. The difference is close to zero for the vast majority of the pixels and the mean difference over the whole granule is very small (e.g., 0.001 for cloud fraction or -19m for cloud top height).

Note that beginning with version 02.00.00, S5P L2_CLOUD OFFL and NRTI will be different because the OFFL will also incorporate VIIRS cloud mask information.

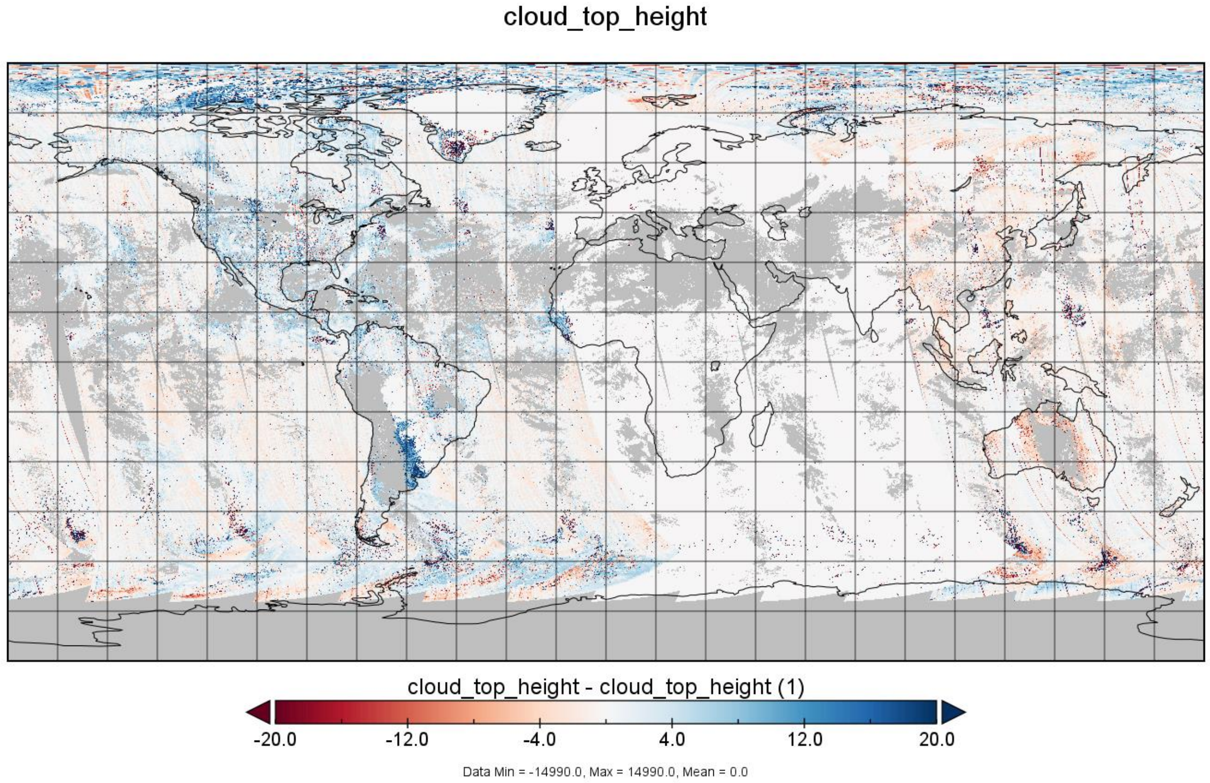


Figure 88: S5P CLOUD CAL 01.01.07 CTH NRTI-OFFL differences on 2019-08-06. CTH expressed in meter.

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 89**. 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

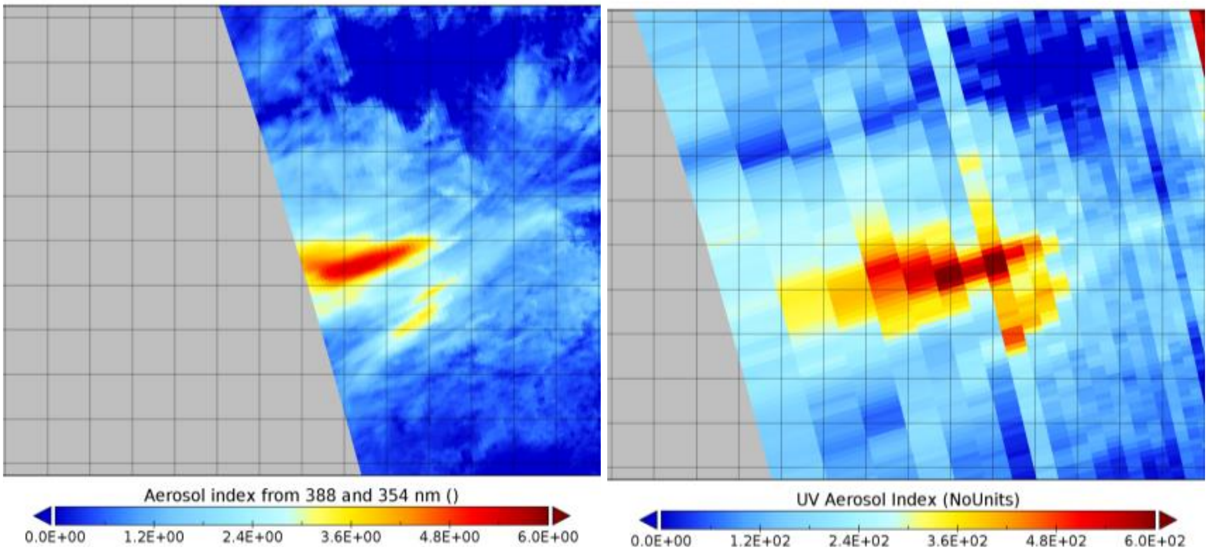
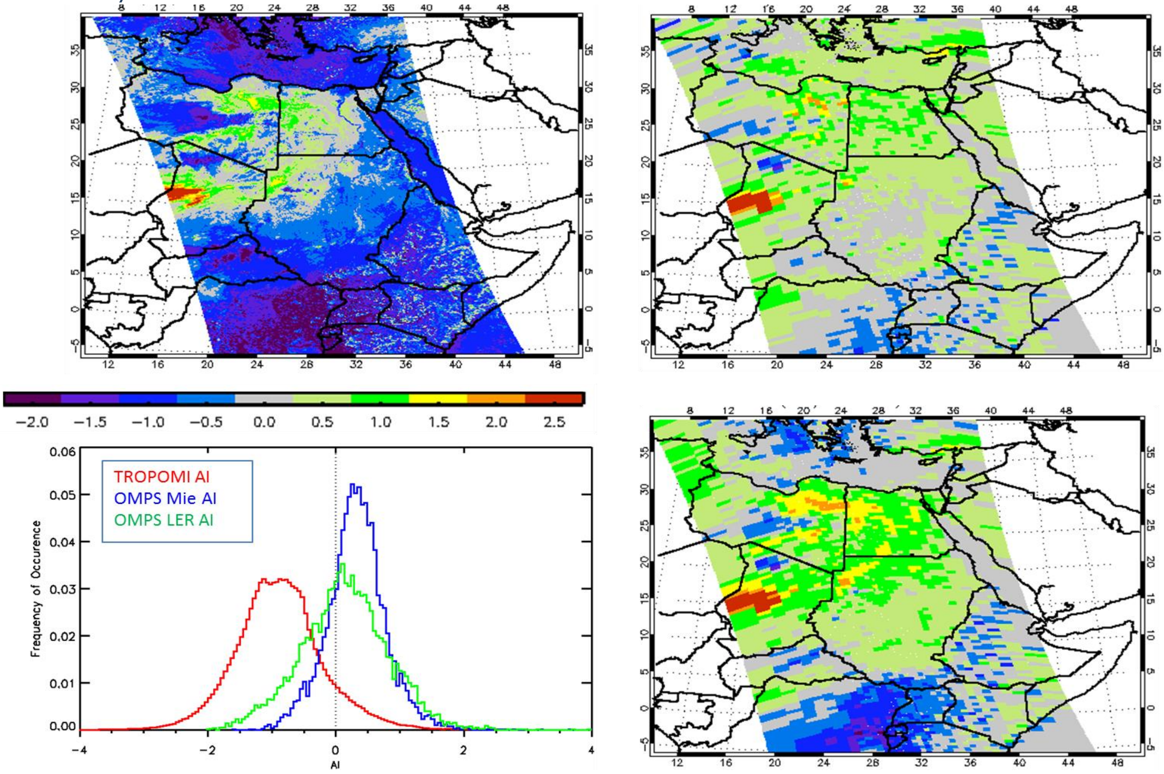


Figure 89: Comparison of S5P TROPOMI UVAI (orbit 00398, left) and OMI OMAERO UV Aerosol Index (orbit 70864, right) for Saharan dust on 10 November 2017. In general very good agreement is found (the stripes in north-south direction in the OMI data are caused by the OMI row anomaly and should be ignored).

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 90** and **Figure 91** below (courtesy of Omar Torres and Changwoo Ahn, NASA-GSFC). At the beginning of TROPOMI measurements (Nov. and Dec. 2017, **Figure 90**), 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 91**). 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 94**). 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

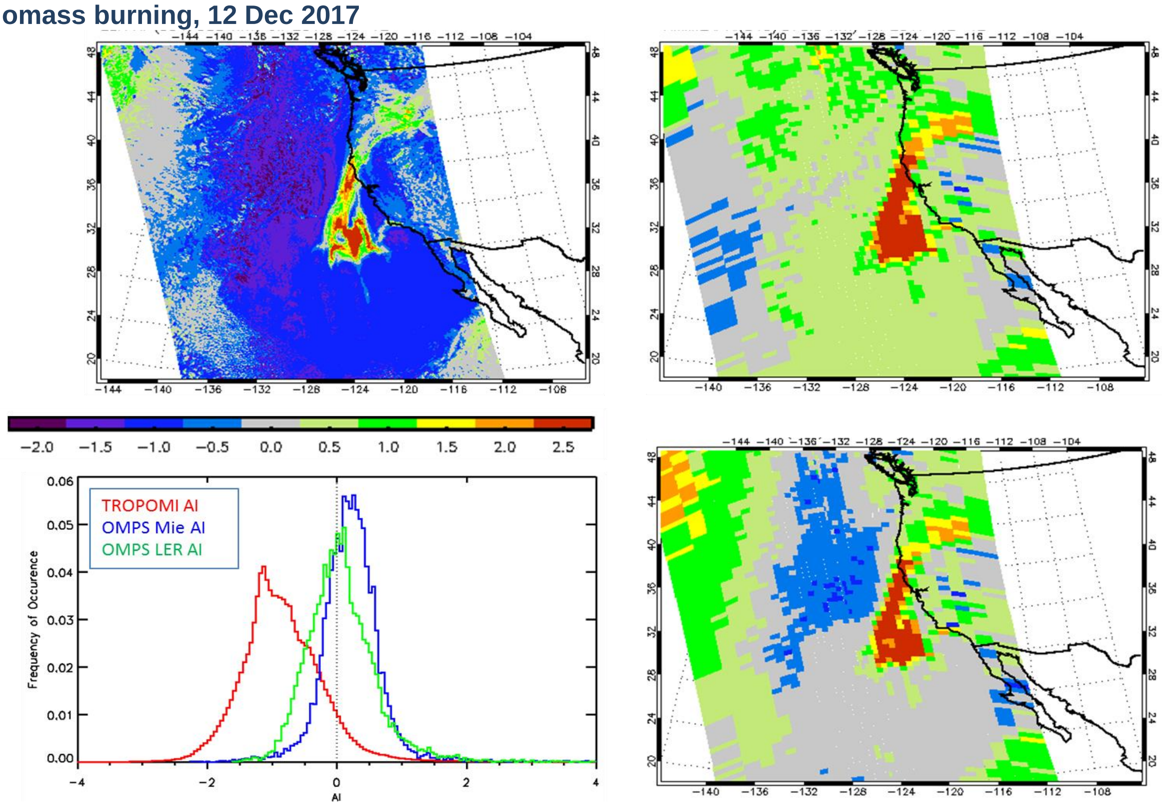


Figure 90: Comparison of UVAI from TROPOMI and OMPs for a situation with desert dust (10 Nov 2017, top) and biomass burning (12 Dec 2017, bottom). For OMPS, UVAI are calculated either assuming LER or Mie clouds. The UVAI for Mie clouds yield more consistent results. The frequency distributions indicate 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).

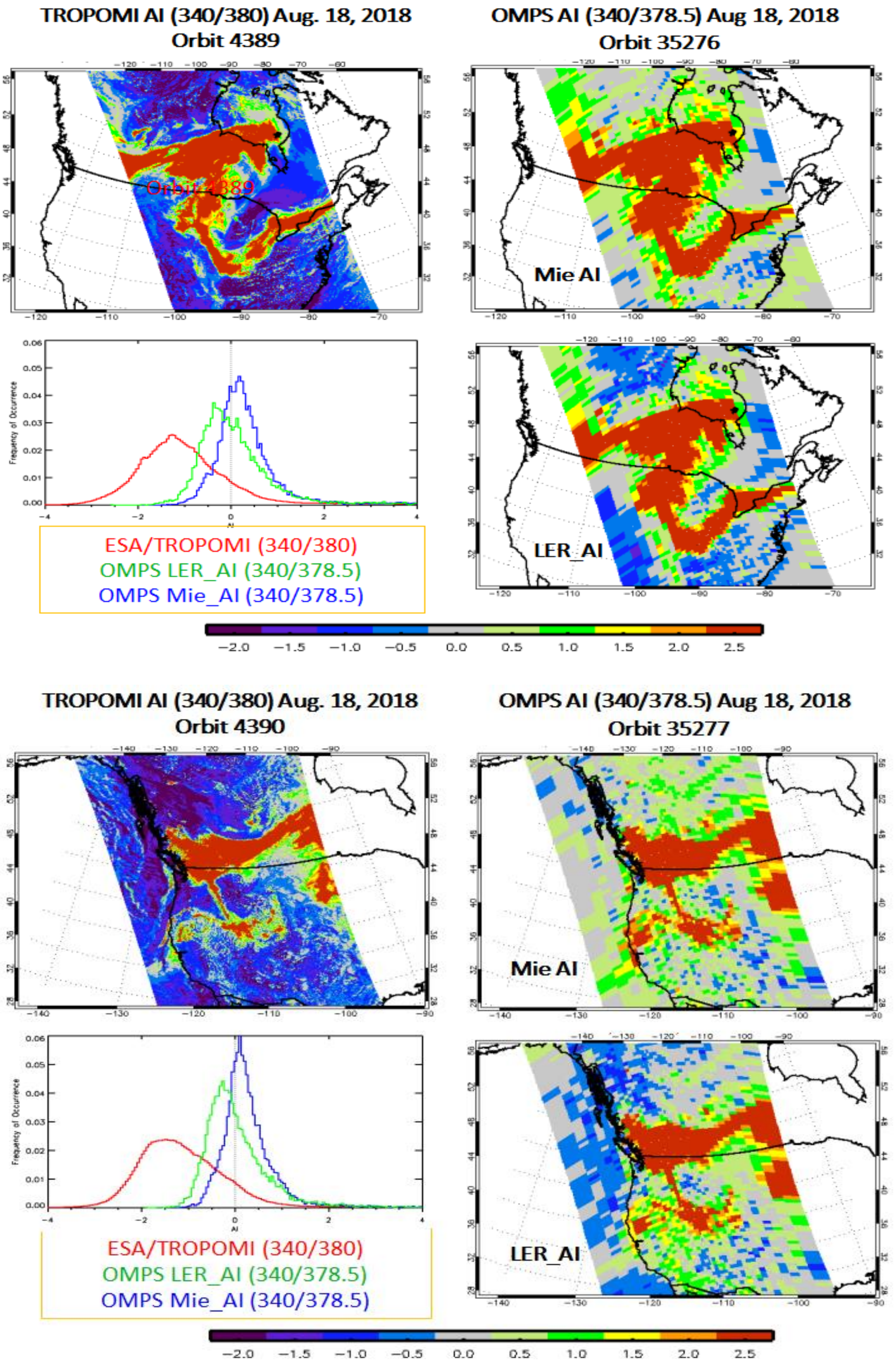


Figure 91: Comparison of UVAI from TROPOMI and OMPs for an observation of a biomass burning plume (18 Aug. 2018). For OMPS UVAI are calculated assuming either LER or Mie clouds. The UVAI for Mie clouds yields more consistent results. In contrast to early TROPOMI observations, values have systematically 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 92** below shows an example of measurements of UVAI and NO_2 VCDs, for which enhanced NO_2 and UVAI are found at the same locations.

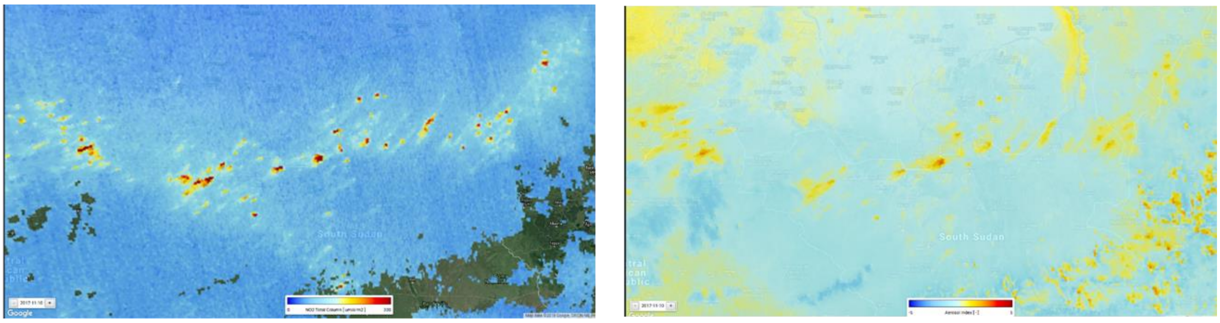


Figure 92: Comparison of NO_2 VCDs (left) and UVAI (right) obtained from S5P TROPOMI for sub-Saharan fires on 10 November 2017.

From the performed validation studies it is concluded that the L2_AER_AI UVAI from S5P TROPOMI is of very good quality and fulfilled the requirements until early 2019. The negative bias found in the S5P TROPOMI data, which continues to increase systematically is outside the bias requirements (+/- 1 UVAI unit) since the beginning of 2019. Here it should be noted that the bias is caused by the degradation of the level 1 irradiance data and will very probably be corrected after the new release of the level 1 data (foreseen at the end of 2019). 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 to February 2020 is shown in **Figure 93** 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 degragation 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.

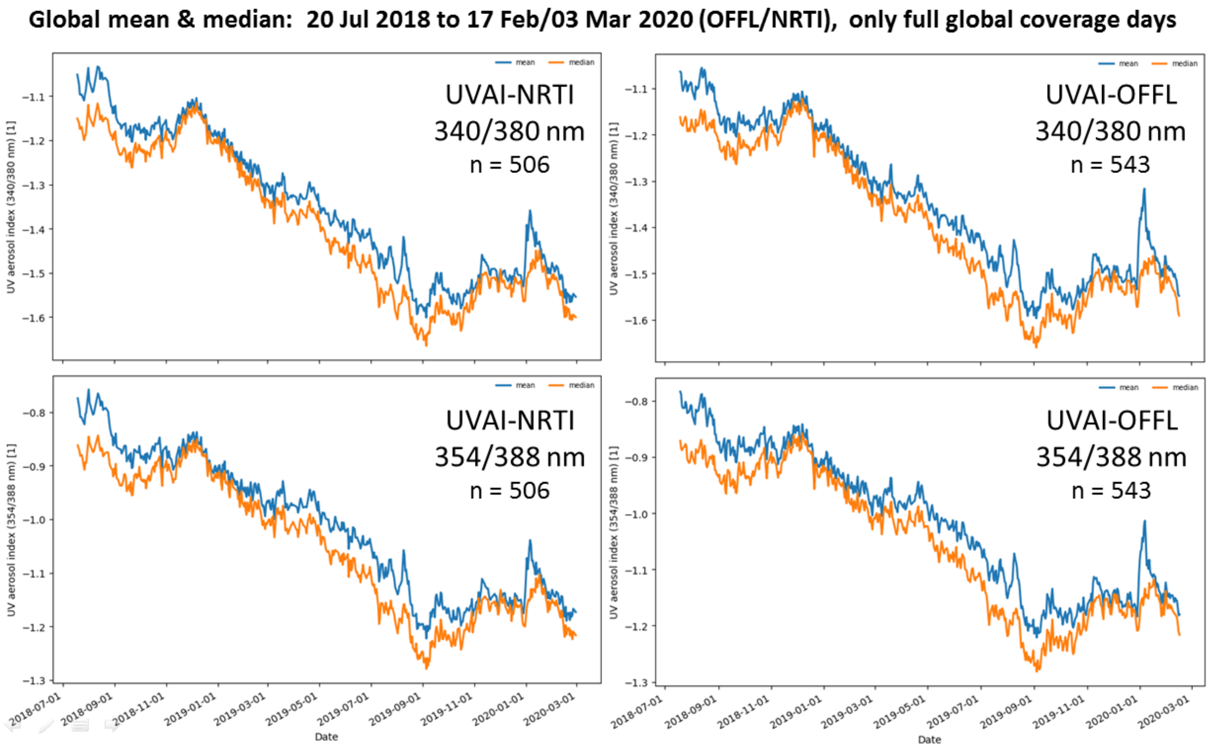


Figure 93: Comparison of the global daily mean (blue) and median (orange) for both L2_AER_AI UVAI wavelength pairs (340/380 and 354/388 nm) and for the NRTI and OFFL data streams, from 20 July 2018 through February 2020.

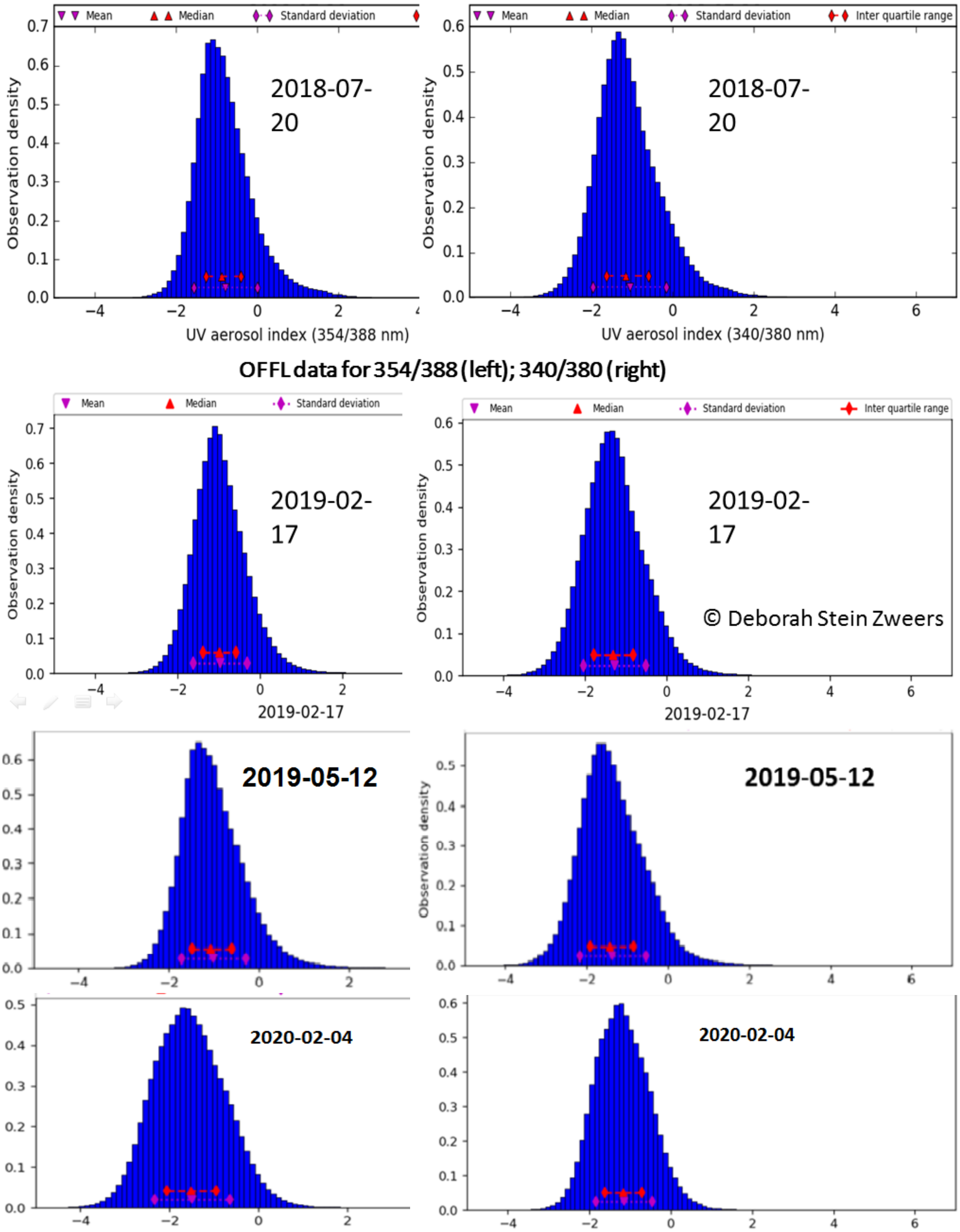


Figure 94: Comparison of the frequency distribution of the UVAI (left: 354/388m,, right: 340/380nm) for four selected days (20 July 2018, 17 February 2019, 19 May 2019, and 04 February 2020)

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.

12.4 Equivalence of L2_AER_AI NRTI and OFFL products

Figure 95 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.

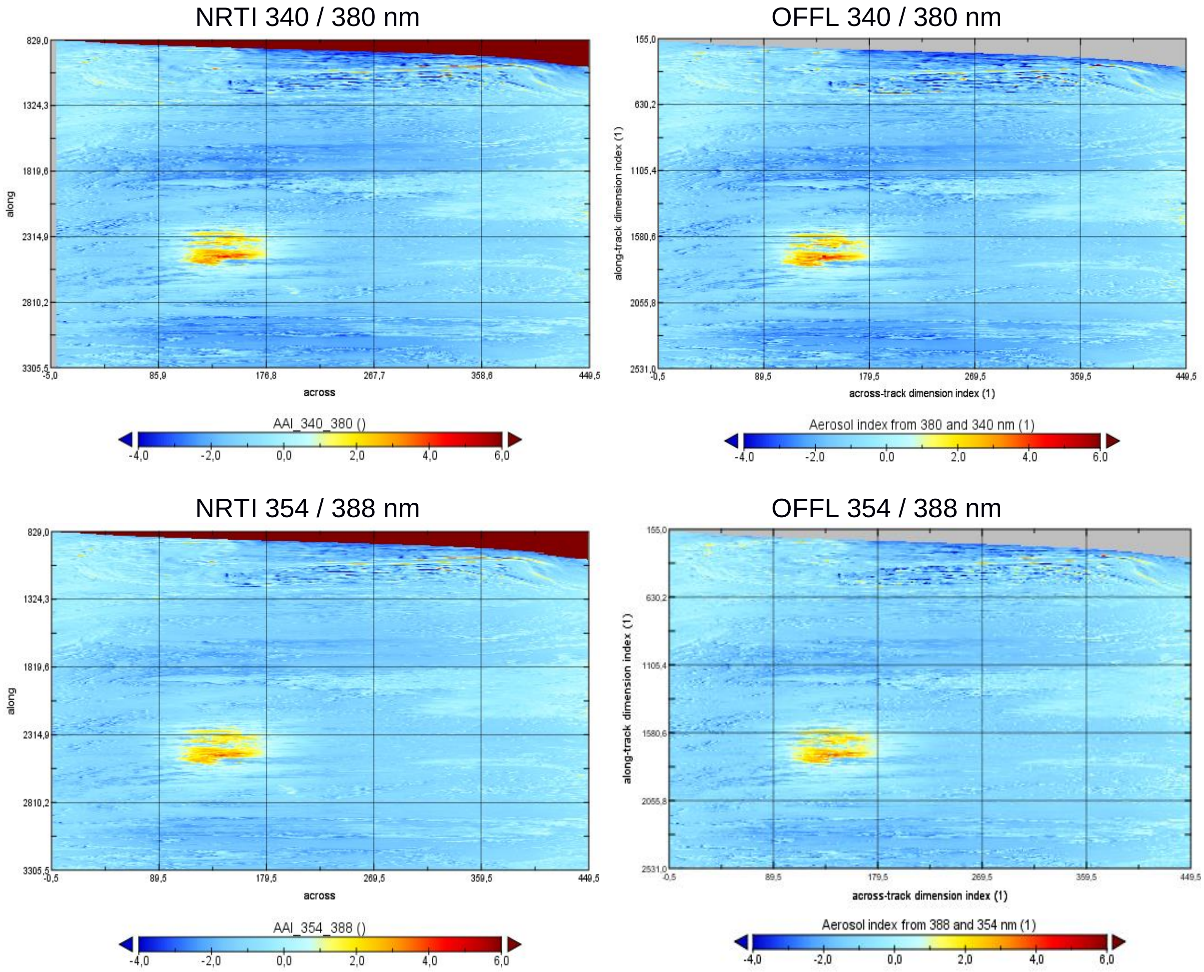


Figure 95: Comparison of the S5P TROPOMI UVAI for a selected orbit (#05033) on 3 October 2018 for the two wavelength pairs (top: 340 / 380 nm, bottom: 354 / 388 nm). While the geographical patterns are the same, the absolute values differ slightly with the NRT values (left) slightly higher than the offline 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 is a new product that 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 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]

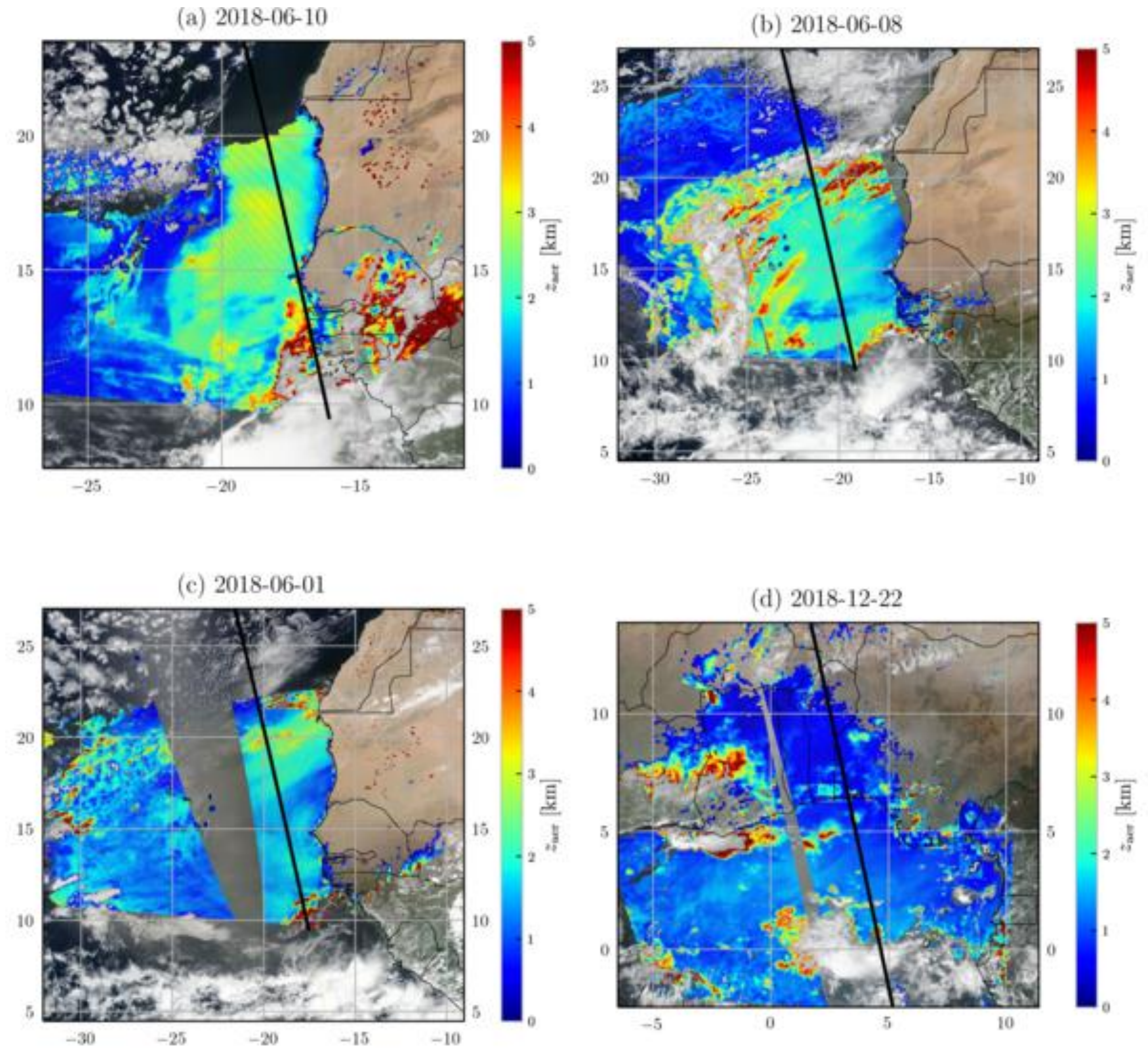


Figure 96: Details of the selected validation cases, showing the ALH on a VIIRS RGB background. The black line represents the Calipso track.

Figure 96 show the cases, (a)-(c) are similar desert dust cases, with dust blowing off the African continent over the Atlantic, and (d) is a smoke case, with smoke over both land and ocean. The black lines display the Caliop tracks, which have a good coverage of the events. The curtain plots from CALIOP are shown in **Figure 97**, 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 center 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 well0-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 ith 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 97 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.

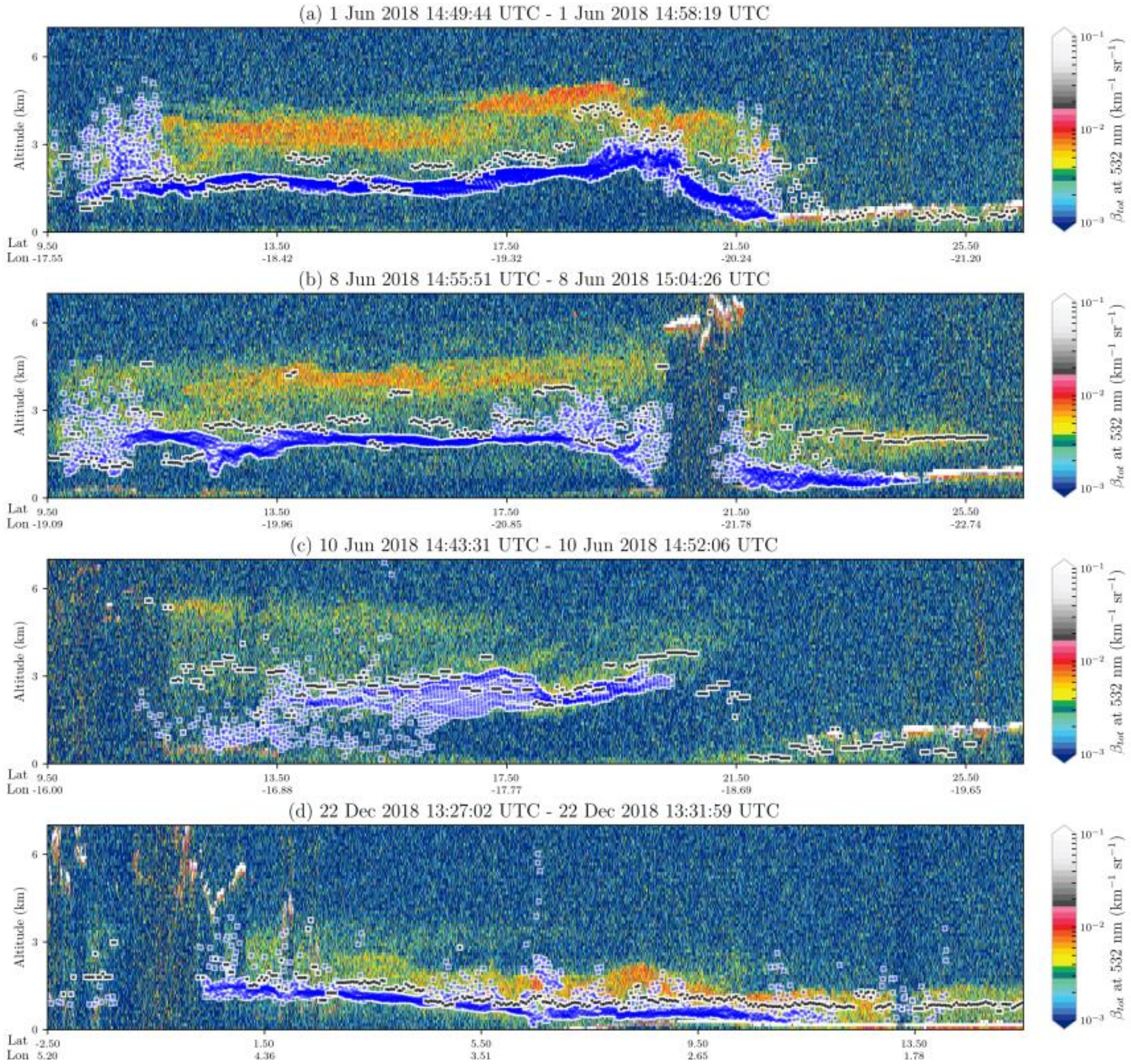


Figure 97: Curtain plots from CALIOP, showing the total attenuated backscatter at 532 nm for the four cases in **Figure 96**. Overplotted on the coloured background image of the total attenuated backscatter are the weighted extinction heights as derived from the backscatter coefficient in black-and-white, and the ALH from collocated TROPOMI pixels in blue-and-white.

In **Figure 98** 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.

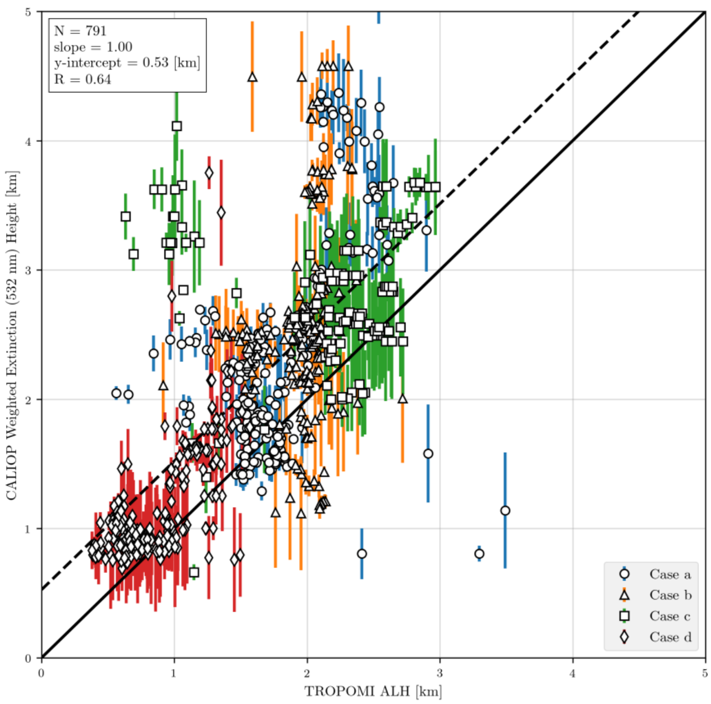


Figure 98: Comparison of ALH from TROPOMI and CALIOP for the cases presented in **Figure 96**. Each case is colour coded. CALIOP weighted extinction height is consistently lower than TROPOMI ALH.

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 99** (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 are have a larger difference, with -2.41 km on average and a median of -1.75 km. The results are very skewed over land, with very large negative values dictating the average — this is indicated by the very large standard deviation of 3.56 km. 50% of the selected colocations over land have an absolute difference with CALIOP weighted extinction height less than approximately 1.0 km. On the right, a similar histogram is shown, but now for only those pixels that have a minimal cost function, or χ^2, smaller than 1E5. The χ^2 represents the goodness-of-fit of the modelled sun-normalised radiances to the observations in the O2-A band, and therefore is a measure of the representativeness of the model (of a simple one aerosol layer atmosphere with known surface reflectance) to reality. Smaller χ^2 indicate a better fit. The retrievals over land generally have

much higher χ^2, and therefore are less reliable. The right panel in **Figure 99** 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.

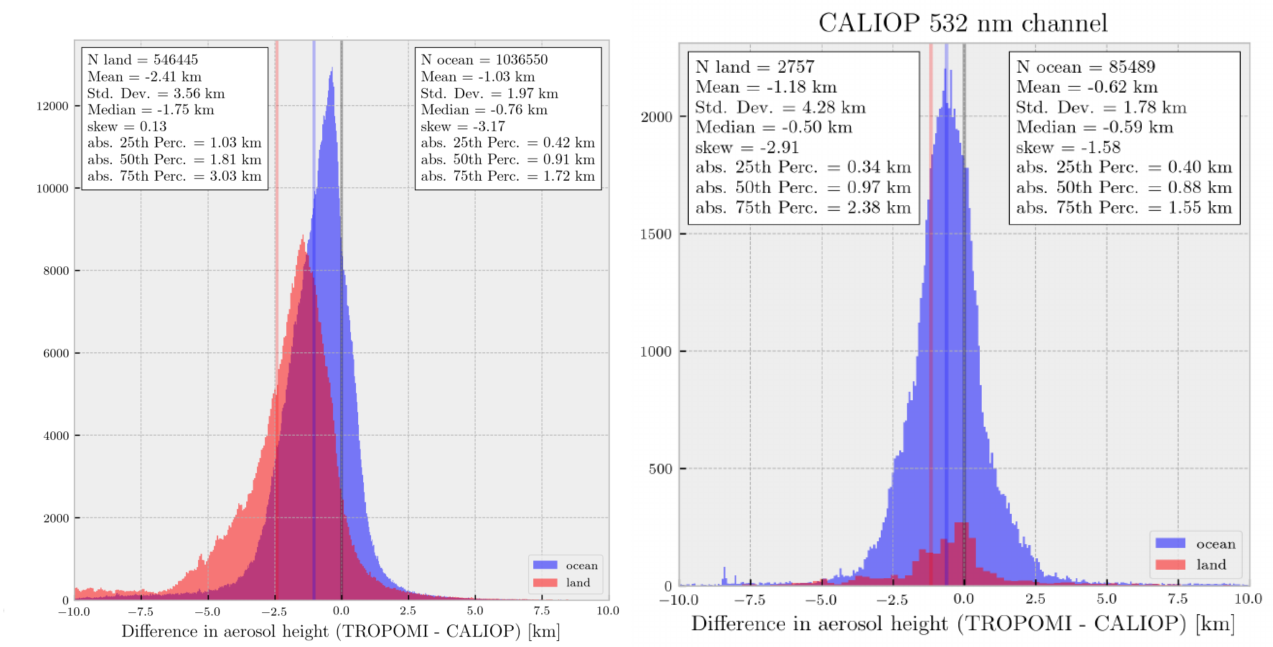


Figure 99: Histogram of differences between CALIOP weighted extinction height and TROPOMI ALH from collocated data between 1 May 2018 and 28 February 2019 (left). The right panel shows the same histogram, but for pixels which were screened for a minimal cost function (chi-squared) smaller than 1E5.

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 100** and **Figure 101**. 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.

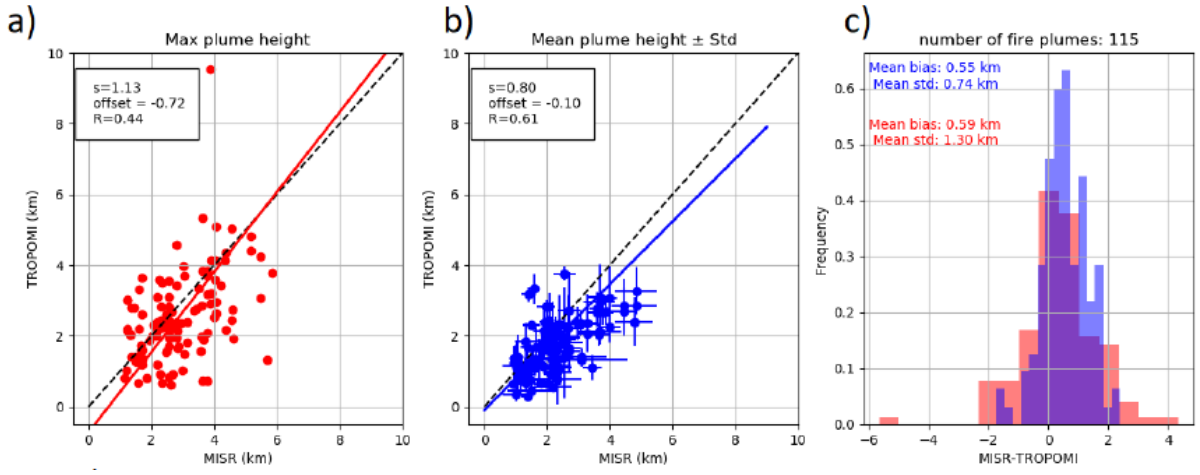


Figure 100: Comparison of TROPOMI ALH and MISR plume height for 115 fires over Nothern America in 2018. See Griffin et al, 2019 for details/

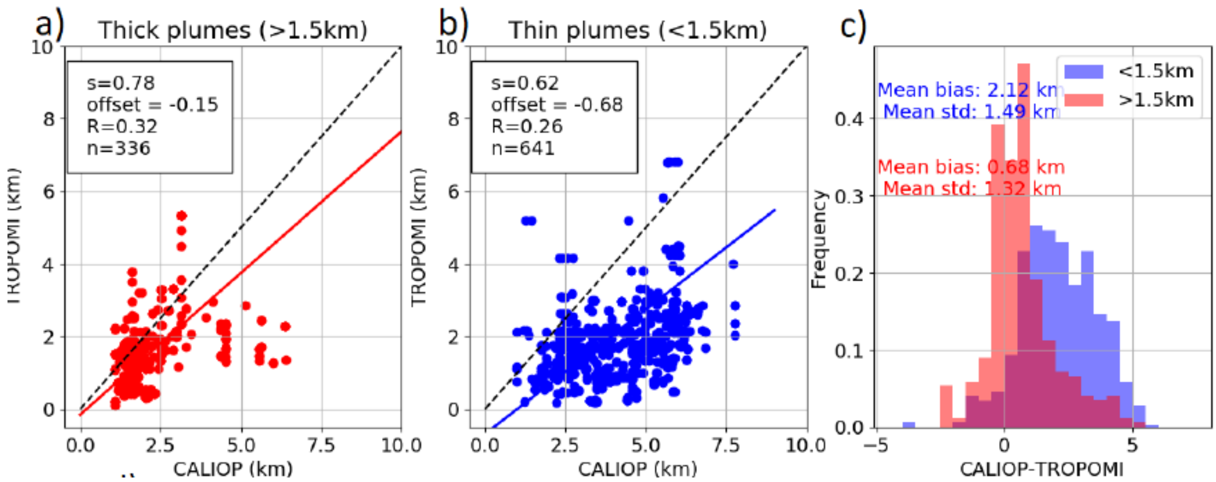


Figure 101: Comparison of TROPOMI ALH and CALIOP average aerosol layer height (top minus bottom of aerosol layer as defined by the feature mask) for collocated pixels near fires over Nothern America in 2018. See Griffin et al, 2019, for details.

13.3.3 Bias

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

13.3.4 Dispersion

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

13.3.5 Dependence on influence quantities

The TROPOMI ALH is strongly dependent on subpixel clouds, and cloud filtering remains essential. The user is strongly encouraged to use all available cloud filters. The ALH is only processed for UV Aerosol Indices larger than zero. However, the UV AI is biased and degrading, which means fewer ALH pixels are processed as time continues. Currently, no significant reduction of the number of pixels can be attributed to degradation of the UV-AI. A small reduction in the number of pixels is observed since August 2019 (see **Figure 102**), but this may be a seasonal effect. 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.

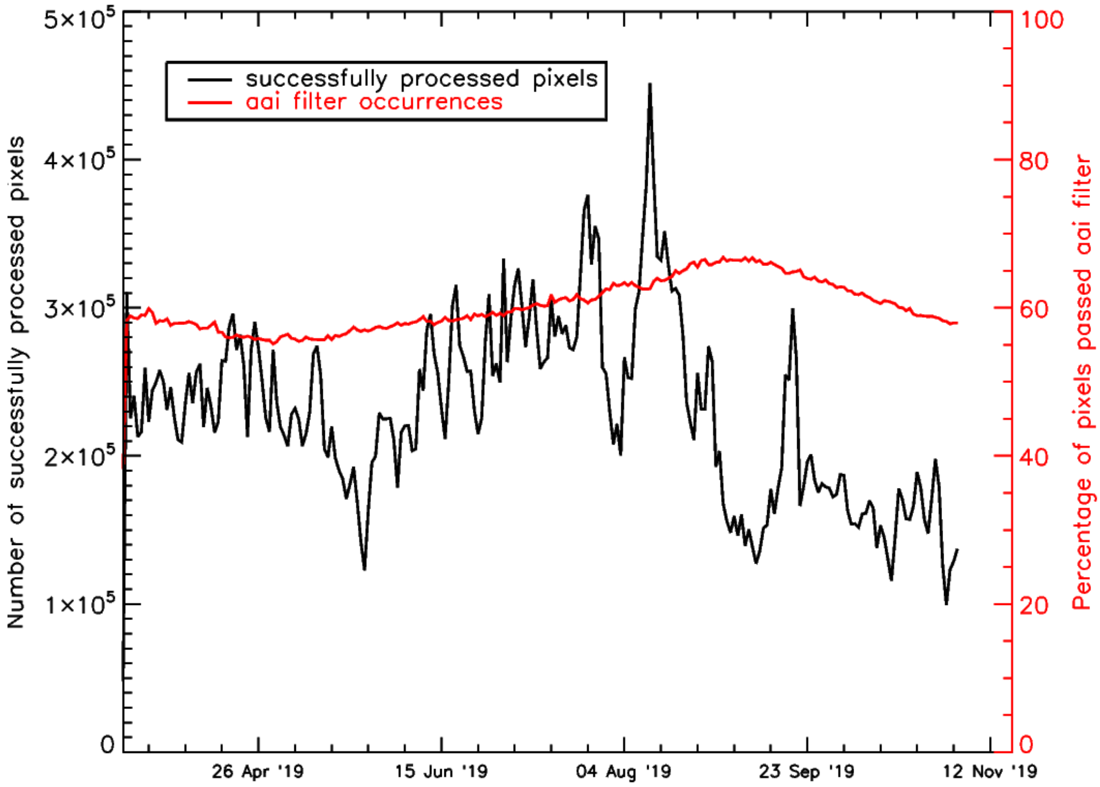


Figure 102: Number of successfully processed ALH pixels (black) and the percentage of pixels that passed the UV-AI > 0 criterion (red) from Apr 2019 to Nov. 2019.

13.3.6 Short term variability

The short term variability of the TROPOMI ALH was not investigated. The products is strongly event-driven, and generaly 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 103** 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.

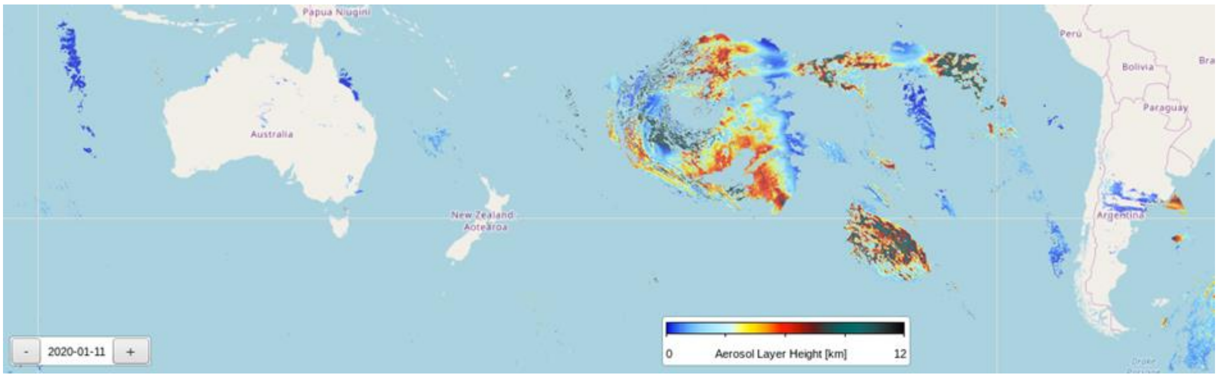


Figure 103: TROPOMI AER_LH product on 11 January 2020 over the south Pacific, showing the altitude as derived by the S5P AER_LH algorithm of the fires smoke from Australian bush fires.

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 104** 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 105** the CALIOP total attenuated backscatter at 532 nm is shown for the yellow track shown in **Figure 104**. The plume can be seen around about 44^{o}S and 110^{o}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.

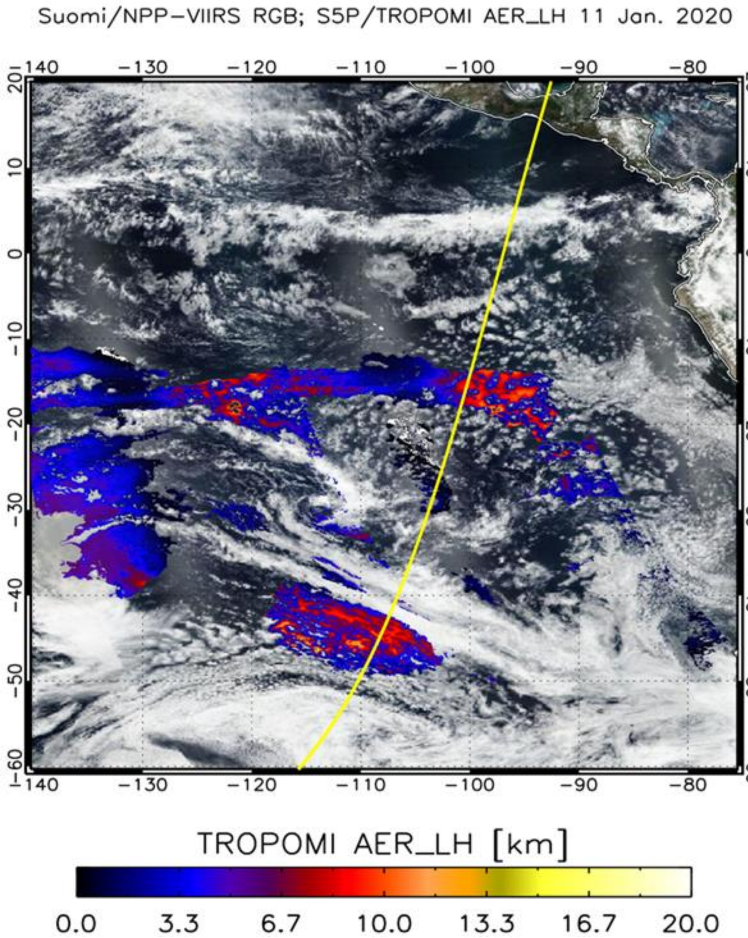


Figure 104: NPP/VIIRS RGB image with S5P/TROPOMI AER_LH on 11 January over the south Pacific with the CALIPSO track of that day overlaid in yellow

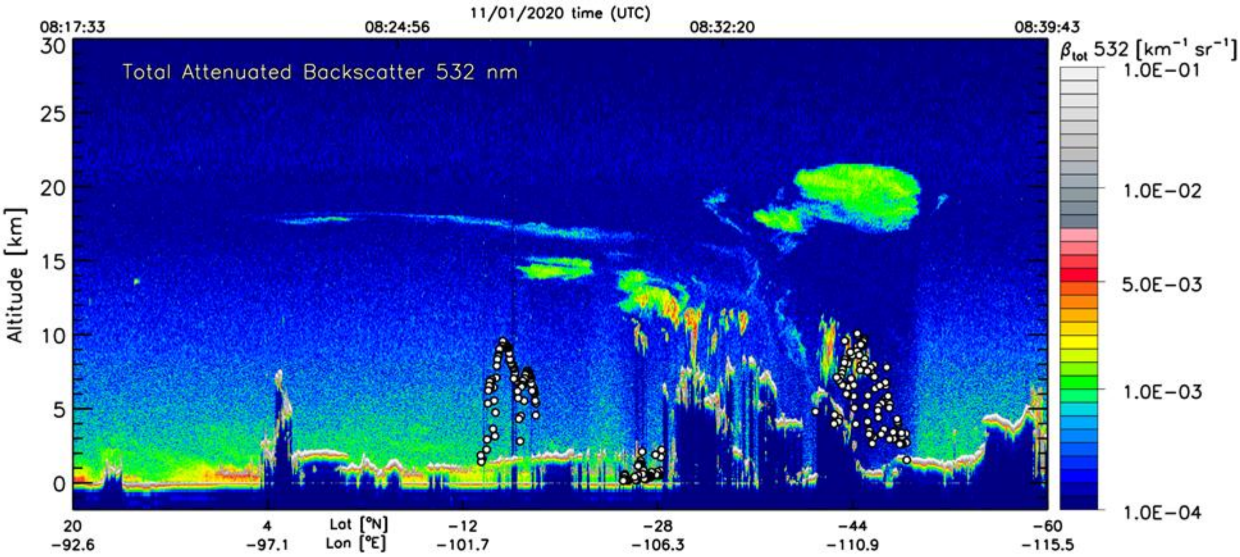


Figure 105: CALIOP aerosol backscatter rat 532 nm along the track shown in **Figure 104** (eastern of Australia on 11 January 11, 2020).

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 104** and **Figure 105**, 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:** 1; **revision:** 0; **date:** 2012-09-20

[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 Reference 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.

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

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

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

Garane, K., Koukouli, M.-E., Verhoelst, T., Fioletov, V., Lerot, C., Heue, K.-P., Bais, A., Balis, D., Bazureau, A., Dehn, A., Goutail, F., Granville, J., Griffin, D., Hubert, D., Keppens, A., Lambert, J.-C., Loyola, D., McLinden, C., Pazmino, A., Pommereau, J.-P., Redondas, A., Romahn, F., Valks, P., Van Roozendael, M., Xu, J., Zehner, C., Zerefos, C., and Zimmer, W.: TROPOMI/S5ptotal ozone column data: global ground-based validation & consistency with other satellite missions, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2019-147, in review, 2019.

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

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

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

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

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

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

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

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

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

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

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

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

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

Nanda, S, De Graaf, M., Veefkind J.P., Sneep M. and Ter Linden, M., Validating TROPOMI aerosol layer height retrievals with CALIOP data, Atmos. Meas. Tech. Disc, 19,https://www.atmos-meas-tech-discuss.net/amt-2019-348/, 2019.

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

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

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.

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.

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.

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/utils/common/documents/jcgm/JCGM_100_2008_E.pdf
[ER_VIM]	International Vocabulary of Metrology (VIM)	http://www.bipm.org/en/publications/guides/vim.html
[ER_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://tccon-wiki.caltech.edu
[ER_TOLnet]	Tropospheric Ozone Lidar Network (TOLnet) website	http://www-air.larc.nasa.gov/missions/TOLNet
[ER_WOUDC]	World Ozone and Ultraviolet Data Centre (WOUDC) website	http://woudc.org

15 Acknowledgements

This Section acknowledges the authors of this report in charge of the MPC Routine Operations validation service (**Table 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).

<table>
<tr><th>S5P Product ID</th><th>Geophysical Quantity</th><th>Product Coordinator for Routine Operations Validation Activities</th><th>Product Contributors to Routine Operations Validation Activities</th></tr>
<tr><td>L1B</td><td>Radiance and irradiance</td><td>Q. Kleipool (KNMI)</td><td></td></tr>
<tr><td>L2_O3</td><td>O_3 total column</td><td>T. Verhoelst (BIRA-IASB)</td><td>K. Garane (AUTH)
K.-P. Heue (DLR)
C. Lerot (BIRA-IASB)</td></tr>
<tr><td>L2_O3_PR</td><td>O_3 profile</td><td>A. Keppens (BIRA-IASB)</td><td>O. Tuinder (KNMI)</td></tr>
<tr><td>L2_O3_TCL</td><td>O_3 tropospheric column</td><td>D. Hubert (BIRA-IASB)</td><td>K.-P. Heue (DLR)</td></tr>
<tr><td rowspan="3">L2_NO2</td><td>NO_2 stratospheric column</td><td rowspan="3">K.-U. Eichmann (IUPB)</td><td>T. Verhoelst (BIRA-IASB)</td></tr>
<tr><td>NO_2 tropospheric column</td><td rowspan="2">S. Compernolle (BIRA-IASB)
H. Eskes (KNMI)
P. Valks (DLR)</td></tr>
<tr><td>NO_2 total column</td></tr>
<tr><td>L2_SO2</td><td>SO_2 total column</td><td>T. Wagner (MPI-C)</td><td>P. Hedelt (DLR)
N. Theys (BIRA-IASB)</td></tr>
<tr><td>L2_HCHO</td><td>HCHO total column</td><td>K.-U. Eichmann (IUPB)</td><td>K.L. Chan (DLR)
S. Compernolle (BIRA-IASB)
I. De Smedt (BIRA-IASB)</td></tr>
<tr><td>L2_CO</td><td>CO total column</td><td>B. Langerock (BIRA-IASB)</td><td>J. Landgraf (SRON)
M.K. Sha (BIRA-IASB)</td></tr>
<tr><td>L2_CH4</td><td>CH_4 total column</td><td>M.K. Sha (BIRA-IASB)</td><td>J. Landgraf (SRON)
B. Langerock (BIRA-IASB)</td></tr>
<tr><td rowspan="3">L2_CLOUD</td><td>Cloud Fraction</td><td rowspan="3">S. Compernolle (BIRA-IASB)</td><td>R. Lutz (DLR)
P. Wang (KNMI)</td></tr>
<tr><td>Cloud Height</td><td>A. Argyrouli (DLR)
S. Compernolle (BIRA-IASB)</td></tr>
<tr><td>Cloud Optical Thickness</td><td>P. Wang (KNMI)</td></tr>
<tr><td>L2_AER_AI</td><td>Aerosol Absorbing Index</td><td rowspan="2">T. Wagner (MPI-C)</td><td>D. Stein Zweers (KNMI)</td></tr>
<tr><td>L2_AER_LH</td><td>Aerosol Layer Height</td><td>M. de Graaf (KNMI)</td></tr>
</table>

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. Compernolle, J. Granville, D. Hubert, A. Keppens, J.-C. Lambert, and T. Verhoelst. Multi-TASTE has been supported by the Belgian Federal Science Policy Office (BELSPO), with additional support provided by the EC, ESA and EUMETSAT in the context of several satellite validation and metrology projects.

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

15.3 Validation data

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

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

- Rapid delivery O_3 column data from the LATMOS_RT facility at IPSL/UVSQ, WMO's Ozone Mapping Centre in Thessaloniki and WOUDC in Toronto was gathered in the framework of the S5PVT AO project VALTOZ (ID #28568, PI D. Balis, AUTH, Co-Is ML. Koukouli, E. Zyrichidou, J.-C. Lambert, T. Verhoelst, J. Granville, A. Pazmino, F. Goutail, J.-P. Pommereau, A. Bazureau, and C. Zerefos).
- Rapid delivery O_3 profile data from the SHADOZ network was organised in the framework of the S5PVT AO project CHEOPS-5p (ID #28587, PIs A. Keppens and J.-C. Lambert, BIRA-IASB, Co-Is D. Balis, D. Hubert, W. Steinbrecht, T. Stavrakou, A. Delcloo, S. Godin-Beekmann, T. Leblanc, R. Stübi, A.M. Thompson, T. Verhoelst, G. Ancellet, and V. Duflot). Rapid delivery ozonesonde profile data were also provided by KNMI (A. Piters, M. Allaart) and NOAA (B.J. Johnson).
- Rapid delivery NO_2 data from NDACC MAX-DOAS and ZSL-DOAS stations was gathered in the framework of the S5PVT AO projects CESAR (ID #28596, PI A. Apituley, KNMI) and NIDFORVAL (ID #28607, PI G. Pinardi, BIRA-IASB). The LATMOS SAOZ_RT team (A. Pazmino, A. Bazureau, F. Goutail, and J.-P. Pommereau) at IPSL/UVSQ/UPMC/CNRS is thanked for the near-real-time processing and delivery of ZSL-DOAS SAOZ data. ESA's FRM programme and LuftBlick/U. Innsbruck (A. Cede, M. Gebetsberger and M. Tiefengraber) are acknowledged for the rapid delivery of total NO_2 data from the Pandonia Global Network (PGN).

- Rapid delivery HCHO data from NDACC FTIR and MAX-DOAS stations was gathered in the framework of the S5PVT AO projects CESAR (ID #28596, PI A. Apituley, KNMI) and NIDFORVAL (ID #28607, Co-PIs G. Pinardi and C. Vigouroux, BIRA-IASB).
- Rapid delivery CO and CH_4 data from TCCON FTIR stations was gathered in the framework of the S5PVT AO project TCCON4S5P (ID #28603, PI M. Kumar Sha, BIRA-IASB).
- Rapid delivery of NDACC data is partly supported by the CAMS-27 data procurement service contracted by ECMWF for the validation of the Copernicus Atmospheric Monitoring Service (CAMS).

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

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

EUMETSAT AC-SAF and DLR are acknowledged for the provision of MetOp-A and MetOp-B GOME-2 ozone and cloud data.

KNMI is acknowledged for the provision of EOS-Aura OMI O_3, NO_2, HCHO and UVAI data. The OMI QA4ECV data records are an outcome of the EC FP7-SPACE-2013-1 project No 607405: Quality Assurance for Essential Climate Variables (QA4ECV).

NASA/GSFC is acknowledged for the provision of (i) Suomi-NPP OMPS radiance, O_3 and UVAI data, (ii) Suomi-NPP VIIRS cloud data obtained with a pre-production code run specifically for limited S5P team analysis, (iii) EOS-Aqua MODIS cloud fraction, cloud top height and cloud optical thickness data, and (iv) MISR and CALIOP aerosol layer height data.

15.4 Agency support

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

16 Terms, definitions and abbreviated terms

16.1 Terms and definitions

accuracy	closeness of agreement between a quantity value obtained by measurement and the true value of the measurand; note that it is not a quantity and it is not given a numerical quantity value [JCGM-VIM]
bias	(1) systematic error of indication of a measuring system [JCGM-VIM] (2) estimate of a systematic measurement error [JCGM-VIM]
error	(1) measured quantity value minus a reference quantity value [JCGM-VIM] (2) difference of quantity value obtained by measurement and true value of the measurand (CEOS/ISO)
influence quantity	quantity that, in a direct measurement, does not affect the quantity that is actually measured, but affects the relation between the indication and the measurement result [JCGM-VIM]
Level 1b data	calibrated, geo-located Earth reflectance and radiance spectra in all spectral bands; solar irradiance data, annotation data and references to used calibration data
Level 2 data	geophysical measurand at the same resolution and geolocation as the Level 1 source data from which it is derived
Level 3 data	data or retrieved geophysical parameters (i.e. derived from Level 1 or 2 data products) mapped on uniform space-time grid scales, usually with some completeness and consistency. Such re-sampling may include averaging, compositing, kriging, use of Kalman filters…
measurand	quantity intended to be measured [JCGM-VIM]
measurement bias	estimate of a systematic measurement error [JCGM-VIM]
measurement error	measured quantity value minus a reference quantity value [JCGM-VIM]
measurement uncertainty	non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used [JCGM-VIM]
precision	closeness of agreement between quantity values obtained by replicate measurements of a quantity on the same or similar object under specified conditions [JCGM-VIM]
random error	component of measurement error that in replicate measurements varies in an unpredictable manner; note that random measurement error equals measurement error minus systematic measurement error [JCGM-VIM]
relative standard uncertainty	standard measurement uncertainty divided by the absolute value of the measured quantity value [JCGM-VIM]
stability	ability of a measuring system to maintain its metrological characteristics constant with time [JCGM-VIM]
systematic error	component of measurement error that in replicate measurements remains constant or varies in a predictable manner [JCGM-VIM]
uncertainty	non-negative parameter that characterizes the dispersion of the quantity values that are being attributed to a measurand, based on the information used [JCGM-VIM]
validation	(1) the process of assessing, by independent means, the quality of the data products derived from the system outputs (CEOS/ISO) (2) verification where the specified requirements are adequate for an intended use [JCGM-VIM]
verification	the provision of objective evidence that a given data product fulfils specified requirements; note that, when applicable, measurement uncertainty should be taken into consideration [JCGM-VIM]

16.2 Acronyms and abbreviations

A(A)I	Aerosol (Absorbing) Index
AC-SAF	Atmospheric Composition Satellite Application Facility
ACTRIS	European Research Infrastructure for the observation of Aerosol, Clouds, and Trace gases
ALC	Automated Lidars and Ceilometers network
AMF	Air Mass Factor
AO	Announcement of Opportunity
ARM	Atmospheric Radiation Measurement program
ATBD	Algorithm Theoretical Basis Document
AVS	Automated Validation Server
AUTH	Aristotle University of Thessaloniki
BELSPO	Belgian Federal Science Policy Office
BIRA-IASB	Royal Belgian Institute for Space Aeronomy
C3S	Copernicus Climate Change Service
CAL	Clouds As Layers
CAMS	Copernicus Atmosphere Monitoring Service
CCD	Convective Cloud Differential method
CCI	Climate Change Initiative
CESAR	Cabauw Experimental Research Site for Atmospheric Research
CF	Cloud Fraction (fractional cloud cover)
CHEOPS-5p	Validation of Copernicus HEight-resolved Ozone data Products from Sentinel-5p
CLOUDNET	Cloud properties monitoring Network
COT	Cloud Optical thickness
CRB	Clouds as Reflecting Boundaries
CRG	Climate Research Group
C(T)H	Cloud (Top) Height
DLR	German Aerospace Center / Deutsches Zentrum für Luft- und Raumfahrt
DOAS	Differential Optical Absorption Spectroscopy
DU	Dobson Unit
EARLINET	European Aerosol Research Lidar Network
EC	European Commission
ECMWF	European Centre for Medium-Range Weather Forecasts
EOS	Earth Observing System
EPS	EUMETSAT Polar System
ESA	European Space Agency
ESL	Expert Support Laboratory
EU	European Union
EUMETNET	European Meteorological Services Network
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites

EVDC	ESA Atmospheric Validation Data Centre
FRM	Fiducial Reference Measurement
FTIR	Fourier Transform Infra-Red
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