S5P Mission Performance Centre Formaldehyde [L2__HCHO___] Readme

document number	S5P-MPC-BIRA-PRF-HCHO	
issue	1.1	
date	2018-12-18	
product version	V01.01.05	
status	Released	
Prepared by	I. De Smedt (BIRA-IASB), K.-U. Eichmann (IUPB)	MPC Product Lead MPC VAL Product Coordinator
Reviewed by	J.-C. Lambert (BIRA-IASB), D. Loyola (DLR), J. P. Veefkind (KNMI), A. Dehn (ESA)	MPC ESL-VAL Lead MPC ESL-L2 Lead MPC Technical Manager MPC Project Officer
Approved by	A. Dehn (ESA), C. Zehner (ESA)	ESA Data Quality Manager ESA Mission Manager

MPC Contributors	Steven Compernolle (BIRA-IASB), Bavo Langerock (BIRA-IASB), K.L Chan (DLR)	MPC ESL-VAL MPC ESL-VAL MPC ESL-L2
S5PVT Contributors	Gaia Pinardi (BIRA-IASB), Corinne Vigouroux (BIRA-IASB)	NIDFORVAL PI (AO ID #28607) NIDFORVAL PI (AO ID #28607)
Signatures	Isabelle De Smedt Digitally signed by Isabelle De Smedt Date: 2018.12.18 16:20:45 +01'00' MPC Product Lead / PRF Lead Editor	
	Angelika Dehn Digitally signed by Angelika Dehn DN: o=European Space Agency, c=FR, cn=Angelika Dehn, email=angelika.dehn@esa.int Date: 2018.12.18 14:03:58 +01'00' Angelika Dehn (ESA), Data Quality Manager	
	 18.12.2018 Claus Zehner (ESA), Sentinel-5 Precursor Mission Manager	

1 Summary

This is the Product Readme File (PRF) for the Copernicus Sentinel 5 Precursor Tropospheric Monitoring Instrument (S5p/TROPOMI) Formaldehyde Level 2 data product and is applicable for the Near Real Time (NRTI) and Offline (OFFL) products.

Product Identifier: **L2__HCHO____**

Example filename:

S5P_NRTI_L2__HCHO___20180901T034828_20180901T035328_04580_01_010102_20180901T044053.nc

S5P_OFFL_L2__HCHO___20180815T134454_20180815T152624_04344_01_010102_20180821T154713.nc

The OFFL product has the following Digital Object Identifier (DOI): **10.5270/S5P-tjlxfd2**

The Readme file describes the current processing baseline, product and quality limitations, and product availability status. More information on this data product is available from the Sentinel product webpage:

https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms,

and from the TROPOMI product webpage http://www.tropomi.eu/data-products.

The data file contains the `formaldehyde_tropospheric_vertical_column` which gives the total atmospheric column between the surface and the tropopause. The random error uncertainty originating from the spectral fit is given in the `formaldehyde_tropospheric_vertical_column_precision`. Other uncertainty terms are provided in the support_data, as for example the systematic error uncertainties, with or without contribution from the *a priori* profiles errors (`formaldehyde_tropospheric_vertical_column_trueness` and `formaldehyde_tropospheric_vertical_column_kernel_trueness`). As a user guideline for the data quality, a `qa_value` is given with the data. In order to avoid misinterpretation of the data quality, it is recommended at the current stage to only use those pixels with a `qa_value` above 0.5.

Note that the HCHO data product may be used in different ways, and different fields in the file are relevant depending on the application. For this, we refer to the product user manual [RD03]. The averaging kernels and the *a priori* profiles are provided and should be used for e.g. comparisons with models or profile measurements.

Data product requirement from the S5p Calibration and Validation Plan [RD01]:

Parameter	Data product	Vertical Resolution	Bias	Random
Formaldehyde	HCHO	Tropospheric column	40-80%	1.2e16 (4e15) molec.cm^{-2}

Table 1: HCHO data product requirement extracted from the S5p Calibration and Validation Plan [RD01]

Independent validation by the S5p Mission Performance Centre (MPC) Validation Data Analysis Facility and the Sentinel-5 Precursor Validation Team (S5PVT) concludes that the TROPOMI formaldehyde column data is in good overall agreement with (i) reference measurements collected from ground-based monitoring networks, and (ii) the corresponding satellite data products from GOME-2 and OMI. The mean bias of roughly -50% found between TROPOMI and reference data is within the mission requirements of maximum 40-80%. The scatter of the difference around this mean bias also complies with mission requirements.

2 Processing baseline description

The history of the HCHO processor versions is detailed in Table 2.

Processor Version	In operation from	In operation until
01.01.02	NRTI: orbit 5003, 2018-10-01	Orbit 5929, 2018-12-05
01.01.05	NRTI: orbit 5932, 2018-12-05	Current version
01.01.05	OFFL: orbit 5932, 2018-12-05	Current version

Table 2: History of HCHO processor versions

3 Product Quality

3.1 Recommendations for data usage

In order to avoid misinterpretation of the data quality, it is recommended to only use those TROPOMI pixels associated with a `qa_value` above 0.5 (no error flag, cloud radiance fraction at 340 nm<0.5, Solar Zenith Angle (SZA)<=70°, surface albedo<=0.2, no snow/ice warning, air mass factor>0.1).

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) [RD03] and Algorithm Theoretical Basis Document (ATBD) [RD02] associated with this data product, available on https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms.

3.2 Validation results

3.2.1 Status of product validation

This section presents a summary of the key validation results obtained by the Validation Data Analysis Facility (VDAF) of the S5p MPC and by the S5PVT. It contains preliminary results reported and discussed at the S5p Second Public Release Validation Workshop (September 28, 2018). Individual contributions to the workshop are archived in https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-5p/calibration-validation-activities/sentinel-5p-second-products-release-workshop, while up-to-date validation results and consolidated validation reports are available through the MPC VDAF website at http://mpc-vdaf.tropomi.eu.

Current conclusions are based on the limited amount of reference measurements available at the time of this first analysis, and on the period covered by the initial S5p dataset. The conclusions summarized hereafter need to be confirmed by a larger amount of 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.

3.2.2 Validation approach

Formaldehyde is a reactive trace gas in the atmosphere. It is photo-oxidized in sunlight to carbon dioxide. The half-life is in the order of less than 50 minutes. It is formed in the troposphere either naturally during the oxidation of hydrocarbons or within engines. While the background concentration is mainly due to methane oxidation (few Pmolec.cm^{-2}), it can reach values that are up to 100 times higher in heavy traffic (~70 Pmolec.cm^{-2}). HCHO exhibits an annual cycle with higher values in the summertime and a significant diurnal cycle, especially in urban environments.

First validation results are based on correlative ground-based measurements from 13 MAXDOAS stations and 18 FTIR stations contributing to NDACC, and on global satellite measurements from OMI and GOME-2. Mainly, OFFL Level 2 products version 1.01.02 data from May to 2018 is used. The bias of the measured total HCHO should be better than 80% and the precision of an individual pixel better than 12 Pmolec.cm^{-2}.

3.2.2.1 Ground-based networks

S5p TROPOMI L2__HCHO__ formaldehyde column data have been compared to reference measurements acquired by MAX-DOAS UV-Visible instruments (Multi-AXis Differential Optical Absorption Spectroscopy) and FTIR instruments contributing to the Network for the Detection of Atmospheric Composition Change (NDACC). Several of these MAXDOAS and FTIR stations have provided data in fast delivery mode through the S5PVT AO project NIDFORVal (ID 28607).

3.2.2.2 Satellites

Initial S5p TROPOMI L2__HCHO__ formaldehyde column data have also been compared to MetOp-A and B GOME-2 data (version GDP 4.8) and to the QA4ECV reprocessing of EOS-Aura OMI data, and to S5p HCHO column data retrieved with an independent IUP retrieval algorithm (University of Bremen).

3.2.3 Validation results

The FTIR-based comparisons at 16 stations providing faster delivery and 2 stations with data only up to May (NIDFORVAL project) are performed with the following criteria: the co-location must be within 30km of the station and 12 hours of the TROPOMI measurement. TROPOMI ground pixels are filtered on the `qa_value` (`qa_value` > 0.5, at least 10 pixels) before calculating a daily mean which will be compared to daily means of the FTIR measurements. The FTIR retrieved profiles are re-gridded to the TROPOMI vertical resolution and smoothed vertically with the TROPOMI averaging kernels. The mean bias between TROPOMI and FTIR data is -10% with a standard deviation of 2.6 Pmolec.cm^{-2} for all sites for the OFFL dataset for all sites and -26% (2.6 Pmolec.cm^{-2}) for the NRT dataset. The bias requirement of 80% is met at all individual stations (except Kiruna), and 40% is reached at 15 stations. The Pearson correlation coefficient is for both datasets better than 0.82. While the standard deviation is within the requirements for clean sites, polluted sites show larger deviations attributed to higher spatial/temporal variability.

The MAXDOAS-based comparisons at 10 stations providing faster delivery (NIDFORVAL project) are done for the period May-September 2018. MAXDOAS systematic error uncertainty is estimated to be 20%, and its precision is 30%. TROPOMI ground pixels are filtered on the `qa_value` (`qa_value` > 0.5), and selected within a radius of 20 km around the station. Both ground-based temporal averaging within +/- 1h and interpolation at TROPOMI overpass time have been tested. The daily data is used if at least 5 ground pixels remain after the colocation selection. The bias requirements of 80% are met at all MAXDOAS stations and 40% is reached at most stations. Outlying results at Mexican sites need further investigation. For polluted sites, the comparison spread is larger than the precision requirement; however, one should be aware that random error of the MAXDOAS data and the comparison mismatch errors also contribute to the comparison spread.

Three MAXDOAS stations have also contributed by delivering data routinely to the VDAF Automated Validation Server, for the period May-September 2018. Only satellite ground pixels intersecting the MAXDOAS measurement area are used, and the closest MAXDOAS measurement in time, within 0.5 h, is kept. The bias requirement of 80% is met. The comparison spread is ~8 Pmolec.cm^{-2}; this is an upper limit to the true satellite precision, as the MAXDOAS random error and comparison mismatch errors also contributes to the comparison spread. This value is below the upper limit of 12 Pmolec.cm^{-2}, therefore this study suggests that the precision requirement is met.

Comparisons of TROPOMI with OMI data exhibit also a negative bias of about 20% that is mainly attributed to differences in the cloud-corrected Air Mass Factors (AMFs).

4 Known Data Quality Issues

Currently, the following data quality issues are known, which are not accounted for by the quality flags, and should be kept in mind when looking at the formaldehyde product itself and also at preliminary validation results.

QA values

The `qa_value` parameters are currently not set correctly over snow/ice regions, above 75° of SZA. They also need to be further checked over cloudy scenes.

Surface albedo climatology

The current surface albedo climatology has a spatial resolution of 0.5° x 0.5° and a time resolution of 1 month. This resolution is known to be too coarse compared to the much higher spatial resolution of S5p TROPOMI ground pixels. Localized signatures of highly varying albedo in inhomogeneous scenes can be seen in the HCHO columns.

Sun glint

For data up to 16 Oct. 2018, the quality of the HCHO columns was reduced because sun glint was present in the equatorial Pacific, i.e. the reference sector used for the background correction. As cloud fraction and cloud albedo are affected by sun glint, the cloud-corrected AMFs present too large values in the reference sector and the HCHO columns are affected. From 17 Oct., sun glint is out of the reference sector and the quality of the HCHO columns is back to normal. An update of the algorithm is in preparation, in order to avoid that issue appears again.

A priori profiles from TM5 model

The NRTI and OFFL processing use TM5 data for the same time period but from slightly different model settings (e.g. meteorological input data). This is expected and can lead to small differences between NRTI and OFFL `formaldehyde_tropospheric_vertical_column` (less than 10% in more than 90% of the cases). Most of the discrepancy between NRTI and OFFL is for the last orbit of the day, due to the calendar day change. This issue is not critical, as it happens over the Pacific (with no significant HCHO sources), and will be solved in the future.

Bands 3-4 and 6 spatial miss-alignment

The band 3-4 (450 pixels per scanline) footprints are not fully aligned with the band 6 (448 pixels per scanline) ones. In the worst case, the misalignment can be in the order of half a ground pixel. The OCRA algorithm retrieves the Cloud Fraction (CF) using radiances from bands 3 and 4. This is an *a priori* to the ROCINN algorithm using band 6 spectra. Over heterogeneous scenes, the miss-registration might have an impact on the data quality. In the current products, a shift of two detector pixels between band 3-4 and band 6 is applied based on initial assessment. Due to the lack of the cloud information, the first two pixels of each scanline cannot be analyzed.

Saturation

Some TROPOMI ground pixels might be affected by detector saturation. Those ground pixels should be flagged and their quality is reflected in the `qa_value`. Nevertheless, this effect has very low effect in the UV region (HCHO fitting interval range).

Metadata values exchanged

The global attributes `geospatial_lon_min` and `geospatial_lon_max` values are exchanged; therefore, the user is advised to switch the values for these fields, making note that the `geospatial_lat_min` and `geospatial_lat_max` values are correct. This is an issue traceable to L1b data (version 01.00.00) and is corrected in the following versions of the Level 1B processor.

Orbit numbering in NRTI and OFFL (solved)

Note that NRTI orbit numbers are set with respect to the downlink orbit while OFFL orbit numbers are set with respect to the equator crossing time. This creates an inconsistency between the NRTI and

OFFL orbit numbers which is removed with the activation of processor version **01.01.05** (December 2018.)

5 Algorithm Change Record

For a detailed description of the L2__HCHO____ algorithm, please refer to the ATBD [RD02].

6 Data Format

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

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

Please note that consecutive data granules of the NRTI product show an overlap of about 12 scan lines.

Details of the data format are provided in the Product User Manual (PUM) [RD03].

6.1 Data format changes

There are no changes to report with respect to the previous PRF.

7 Product Availability

The S5p HCHO data are available at https://scihub.copernicus.eu.

More information on this data product and data handling tools are available from the product web page under heading 'Tools': http://www.tropomi.eu/data-products.

For further questions regarding S5p/TROPOMI data products please contact EOSupport@Copernicus.esa.int.

The access and use of any Copernicus Sentinel data available through the Copernicus Sentinel Data Hub is governed by the Legal Notice on the use of Copernicus Sentinel Data and Service Information and is given here:

https://sentinels.copernicus.eu/documents/247904/690755/Sentinel_Data_Legal_Notice.

8 References

[RD01] Sentinel-5 Precursor Calibration and Validation Plan for the Operational Phase
source: ESA; **ref**: ESA-EOPG-CSCOP-PL-0073;
url: https://sentinel.esa.int/documents/247904/2474724/Sentinel-5P-Calibration-and-Validation-Plan.pdf

[RD02] Sentinel-5 precursor/TROPOMI Level 2 Algorithm Theoretical Basis Document Formaldehyde
source: BIRA; **ref**: S5P- BIRA-L2- ATBD-400F;
url: http://www.tropomi.eu/documents/atbd

[RD03] Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Formaldehyde HCHO
source: DLR; **ref**: S5P-L2-DLR-PUM-400F;
url: http://www.tropomi.eu/documents/pum

More information on this data product is available from the Sentinel product webpage:

https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms,

and from the corresponding TROPOMI product webpage http://www.tropomi.eu/data-products.

Abbreviations and acronyms

AMF	Air Mass Factor
ATBD	Algorithm Theoretical Basis Document
AVS	Automated Validation Server
BIRA-IASB	Royal Belgian Institute for Space Aeronomy
CF	Cloud Fraction (fractional cloud cover)
DLR	German Aerospace Center / Deutsches Zentrum für Luft- und Raumfahrt
DOAS	Differential Optical Absorption Spectroscopy
DOI	Digital Object Identifier
ESA	European Space Agency
ESL	Expert Support Laboratory
FRM	Fiducial Reference Measurement
FTIR	Fourier Transform Infra-Red
GOME(-2)	Global Ozone Monitoring Experiment(-2)
IUP-UB	Institute of Environmental Physics – University of Bremen
KNMI	Royal Netherlands Meteorological Institute / Koninklijk Nederlands Meteorologisch Instituut
MAX-DOAS	Multi Axis Differential Optical Absorption Spectroscopy
MetOp	polar orbiting Meteorological Operational satellite
MPC	Mission Performance Centre
NDACC	Network for the Detection of Atmospheric Composition Change
NIDFORVAL	Nitrogen Dioxide and Formaldehyde Validation of TROPOMI
OMI	Ozone Monitoring Instrument
OMPS	Ozone Mapper and Profiling Suite
PRF	Product Readme File
PUM	Product User Manual
S5P	Sentinel-5 Precursor
S5PVT	Sentinel-5 Precursor Validation Team
Suomi NPP	Suomi National Polar-orbiting Partnership
SZA	Solar Zenith Angle
TROPOMI	Tropospheric Monitoring Instrument
VDAF	Validation Data Analysis Facility