

S5P Mission Performance Centre Formaldehyde [L2__HCHO__] Readme



document number	S5P-MPC-BIRA-PRF-HCHO	
issue	2.7	
date	2023-11-29	
product version	V02.06.01	
status	Released	
Prepared by	I. De Smedt (BIRA-IASB) Fabian Romahn (DLR) KU. Eichmann (IUPB)	MPC Product Lead MPC ESL-L2 processor lead MPC VAL Product Coordinator
Reviewed by	JC. Lambert (BIRA-IASB) D. Loyola (DLR) D. Stein Zweers (KNMI)	MPC ESL-VAL Lead MPC ESL-L2 Lead MPC Technical Manager
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) Lidia Saavedra de Miguel (ESA/Serco) M. Alparone (ESA/Serco)	MPC ESL-VAL MPC ESL-VAL MPC ESL-L2 ESA S5p mission support ESA S5p mission support
S5PVT Contributors	Gaia Pinardi (BIRA-IASB) Corinne Vigouroux (BIRA-IASB)	NIDFORVAL PI (AO ID #28607) NIDFORVAL PI (AO ID #28607)
Signatures	Angelika Dehn (ESA), Sentinel-5 Precursor Data Quality Manager Claus Zehner (ESA), Sentinel-5 Precursor Mission Manager	

CHANGE LOG

Reason for change	Issue	Revision	Date
Table 2: addition of version 01.01.08	1	4	19/03/2020
Updates for processor version 02.01.03	2	0	16/07/2020
Table 2: addition of version 02.01.04	2	1	26/11/2020
 Table 2: adapting to version 02.02.01 of the processor Section 4.1 & section 4.2: some text moved from section 4.1 (Known Data Quality Issues) to section 4.2: (Solved Data Quality Issues) Section 6.1: added format changes related to version 02.02.01 	2	2	05/07/2021
 Table 2: addition of version 02.03.00 Section 6.1: added format changes related to version 02.03.00 Section 4: Added description of the Slant Column Density (SCD) background (BG) correction Section 7: addition of link to major S5p mission data gaps 	2	3	09/03/2022
 Table 2: addition of version 02.04.01 Section 4.2: addition of a "Solved Data Quality Issue" related to version 02.04.01 Section 6.1: added format changes related to version 02.04.01 	2	4	20/07/2022
 Table 2: addition of reprocessed data version 02.04.01 Section 7: updates with information related to gaps on the reprocessed dataset. 	2	5	05/04/2023
Table 2: addition of version 02.05.00	2	6	20/07/2023
 Table 2: addition of version 02.06.01 Table 2: addition of a note on REPRO orbit for version 02.05.00 Section 7: Replaced Open Data Hub with CDSE 	2	7	29/11/2023

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), Offline (OFFL) and reprocessed (RPRO) products.

Product Identifier: L2_HCHO___

Example filename:

S5P_NRTI_L2_HCHO__20190729T000912_20190729T001412_09273_01_020103_20200608T142937.nc

 ${\tt S5P_OFFL_L2_HCHO__20190729T081452_20190729T095819_09278_01_020103_20200520T132844.nc}$

S5P_RPRO_L2_HCHO__20201007T202447_20201007T220617_15471_03_020401_20230126T064501.nc

The OFFL and RPRO products have the following Digital Object Identifier (DOI): <u>https://doi.org/10.5270/S5P-vg1i7t0</u>

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 <u>http://www.tropomi.eu/data-products</u>.

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 priori profiles а errors (formaldehyde tropospheric vertical column trueness and formaldehyde tropospheric vertical column kernel trueness). As a user guideline for the data quality, a ga 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 ga 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.

Note: Starting from processor version 2.4.1, new improved Level 1b version 2.1.0 data products are used as input [RD04].

Independent validation by the S5p Mission Performance Centre (MPC) Cal/Val experts and the Sentinel-5 Precursor Validation Team (S5PVT) conclude that the TROPOMI formaldehyde column data is compliant with the requirements as defined in the **S5p Calibration and Validation Plan** [RD01], see Table 1.

Up to date validation results are available in the Routine Operations Consolidated Validation Reports (ROCVR) that are accessible through the MPC Validation Data Analysis Facility (VDAF) website at http://mpc-vdaf.tropomi.eu. The ROCVR reports are issued quarterly.

Parameter	Data product	Vertical Resolution	Bias	Random
Formaldehyde	НСНО	Tropospheric column	80%	1.2e16 (4e15) molec.cm ⁻²

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

2 **Processing baseline description**

The history of the HCHO processor versions is detailed in Table 2. Note that the processor version for HCHO is changing when there is a change to any of the products belonging to the UPAS processor suite (SO2, HCHO, O3 NRTI, O3 OFFL, Tropospheric O3, CLOUD) even if the change is not affecting the HCHO product.

Processor Version	In operation from	In operation until	Relevant improvements
02.06.01	NRTI: orbit 31750, 2023-11-29 OFFL: orbit 31705, 2023-11-26	Current version	No changes with respect to previous version
02.05.00	NRTI: orbit 29878, 2023-07-20 OFFL: orbit 29818, 2023-07-16 RPRO: orbit 25163, 2022-08-21	Orbit 31750, 2023-11-29 Orbit 31704, 2023-11-26 Orbit 25276, 2022-08-30	 No changes with respect to previous version RPRO data: Users are recommended to use the RPRO data in the period 21-30 August 2022 because those products were generated with an optimal configuration of the auxiliary files (unlike the OFFL data).
02.04.01	OFFL: orbit 24655, 2022-07-17 NRTI: orbit 24697, 2022-07-20 RPRO: orbit 2918, 2018-05-07	Orbit 29817, 2023-07-16 Orbit 29861, 2023-07-19 Orbit 24898, 2022-08-03	 No changes in the HCHO retrievals CLOUD input product: the OCRA L1b soft correction is switched off when the latest L1b version 2.1.0 (with both radiance and irradiance degradation corrected) is processed The variable surface temperature is now written correctly (it was set to fill-value in the previous version) Note 1: It is recommended to use the RPRO products in the orbit range 24655 - 24898, period for which also OFFL products are available. This, because the OFFL products in that range will be removed. Note 2: Starting from this processor version, new improved Level 1b version 2.1.0 data products are used as input [RD04]
02.03.00	OFFL: orbit 22768, 2022-03-06 NRTI: orbit 22813, 2022-03-09	Orbit 24654, 2022-07-17 Orbit 24697, 2022-07-20	 Improve robustness of Background calculation The variables satellite_altitude, satellite_latitude, satellite_longitude,

			<pre>satellite_orbit_phase are now written correctly (were set to fill-value in the previous versions)</pre>
02.02.01	OFFL: orbit 19258, 2021-07-01 NRTI: orbit 19308, 2021-07-05	Orbit 22767, 2022-03-06 Orbit 22811, 2022-03-09	 From this version, the variable surface_altitude_precision is correctly written (previous versions reported fill-value) Added surface temperature and DOAS polynomial coefficients Background calculation: improved robustness of earthshine reference calculation to avoid failures due to the presence of fill values in the input L1b product
			version 2.0.0 data products are used as input [RD04]
02.01.04	OFFL: orbit 16213, 2020-11-29 NRTI: orbit 16259, 2020-12-02	Orbit 19257, 2021-07-01 Orbit 19306, 2021-07-05	No changes with respect to previous version
02.01.03	OFFL: orbit14239, 2020-07-13 NRTI: orbit 14285, 2020-07-16	Orbit 16256, 2020-12-02 Orbit 16212, 2020-11-29	 New surface albedo retrieval algorithm (GE_LER) from TROPOMI for the CLOUD product replaces the climatologies New OCRA cloud-free maps based on TROPOMI instead of OMI (affects CLOUD input product) New cloud flags have been introduced (e.g. ice-clouds) The required interpolation of cloud properties co-registration between band 3-4 and band 6 due to the instrument co-registration issues has been improved in the CLOUD product Improved background correction for HCHO product Updated metadata generation to reflect the improved spatial resolution after 6th Aug 2019 Improved handling of the ECMWF information, reading and deriving snow-ice information and propagating wind-information in the level 2 products Improved the parameter qa_value determination Invalid values of geolocation_flags set to correct values New variables added
01.01.08	OFFL: orbit 12432, 2020-03-07 NRTI: orbit 12482, 2020-03-11	Orbit14238, 2020-07-12 Orbit 14285, 2020-07-16	No changes with respect to previous version
01.01.07	OFFL: orbit 7907, 2019-04-23 NRTI: orbit 8000, 2019-04-30	Orbit 12431, 2020-03-07 Orbit 12482, 2020-03-11	No changes with respect to previous version

01.01.06	OFFL: orbit 7542, 2019-03-28 NRTI: orbit 7632, 2019-04-04	Orbit 7906, 2019-04-23 Orbit 7999, 2019-04-30	 Surface classification climatology updated Fixed a bug in the interpolation of the surface albedo climatology Fixed a problem regarding the retrieved CLOUD product parameters being too close to the <i>a-priori</i> values. This might have affected the calculation of the HCHO in cloudy cases (see section 4.2)
01.01.05	RPRO: orbit 3017, 2018-05-14 OFFL: orbit 5833, 2018-11-28 NRTI: orbit 5932, 2018-12-05	Orbit 5832, 2018-11-28 Orbit 7541, 2019-03-28 Orbit 7631, 2019-04-04	Alignment of the configuration for NRTI, OFFL and RPRO chains regarding the Chemistry Transport Model input, leading to the same product quality (see section 4.2)
01.01.02	NRTI: orbit 5003, 2018-10-01	Orbit 5929, 2018-12-05	Initial operational 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 <u>https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms</u>.

3.2 Validation results

Independent validation by the S5p MPC Cal/Val experts and the S5PVT concludes that the TROPOMI formaldehyde column 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 low bias of S5p with respect to MAX-DOAS measurements is roughly -40% for direct comparisons, and -20% when the S5p and MAX-DOAS averaging kernels are applied mutually. Over 4 particular sites in Europe the median bias is -14.5% with a scatter of 31%. S5p also exhibits a similarly low bias at 27 NDACC FTIR stations with high HCHO columns (-29% for HCHO > 8 Pmolec/cm₂), while a positive bias is observed over clean FTIR sites (+20% for HCHO < 2.5 Pmolec/cm₂). Those bias estimates are within the mission requirements (bias below 80%). The scatter of the difference around this mean bias also complies with mission requirements.

Up to date validation results are available in the ROCVR reports that are accessible through the MPC VDAF website at http://mpc-vdaf.tropomi.eu. The reports are issued quarterly.

4 Data Quality Remarks

4.1 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.

Surface albedo climatology

The current surface albedo climatology has a spatial resolution of $0.5^{\circ} \times 0.5^{\circ}$ 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.

Slant Column Density (SCD) background (BG) correction and reference spectra selection

The retrieved SCDs are corrected for possible bias by a so-called background correction (BG). The SCD correction is calculated by linear interpolation of correction matrices which are updated every day and are based on measurements performed over the last four days (moving averages). In case of data gaps or interruptions, the BG matrices can be insufficiently populated and the correction can be uncertain or no correction is applied at all.

The reference radiance spectra selection used for the spectral fitting follows the same logic as BG. If no reference spectrum can be found, a fallback radiance spectrum is used instead. In this case the global file attribute Status_reference_spectrum is set to `fallback_earth' (instead of `nominal_earth'). This can cause inconsistencies in the data and in the calculation of BG matrices which in turn affects the accuracy of the data (bias), but it is limited to one day only.

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) footprints. In the worst case, the misalignment can be of the order of half a ground pixel. The OCRA algorithm retrieves the Cloud Fraction at Bands 3 and 4 and interpolates it linearly, according to the covered area, to band 6. This is an *a priori* to the ROCINN algorithm which works in band 6. Over heterogeneous scenes this miss-registration might have a large impact on the data quality. The cloud height and optical thickness retrieved in band 6 are interpolated back to the band 3 footprints. Due to missing overlap with the band 6 footprints, the first pixel in band 3 (no overlap) does not contain cloud data and the second pixel in band 3 (only partial overlap), contains cloud products with reduced quality. This is also reflected in the cloud data qa_value .

4.2 Solved Data Quality Issues

A priori profiles from TM5 model (solved in version 01.01.05)

The NRTI and OFFL processing use TM5 data covering 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 was not critical, as it happened over the Pacific (with no significant sources), and was solved with version 01.01.05 (see Table 2).

Orbit numbering in NRTI and OFFL (solved in version 01.01.05)

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 (see Table 2).

Cloud product (solved in version 01.01.06)

The HCHO AMF calculation is based on the CLOUD level 2 data. Therefore, quality issues on CLOUD datasets affect the HCHO data. The HCHO retrieval use the Cloud as a Reflecting Boundary (CRB) cloud model. In some rare cases the CRB cloud algorithm (in the versions 01.01.05 and before) converges towards the a-priori cloud top height of 3.8 km. This might affect the calculation of the HCHO column in cloudy cases. The issue was solved since version 01.01.06 of the CLOUD product (see Table 2).

Metadata/Attributes (solved in version 02.01.03)

The spatial resolution of the TROPOMI measurements is improved by bringing the along track ground pixel size from 7.0 to 5.5 km starting on 6th August 2019. After this operations change, the metadata/Attributes fields related to the spatial resolution, remained **unchanged** (hence not aligned to the improved resolution) – but with version 02.01.03 this is corrected.

QA values (solved in version 02.01.03)

The <code>qa_value</code> parameters are set correctly since version 02.01.03 over snow/ice regions, above 75° of SZA.

Sun glint (solved in version 02.01.03)

For data between 27 September 2018 – 16 October 2018 and between 15 September 2019 – 02 November 2019, 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.

Flagging of Saturation (solved)

Some TROPOMI ground pixels might be affected by detector saturation. Those ground pixels are flagged and their quality is reflected in the qa_value . Nevertheless, in the vicinity of saturated pixels there might be pixels also affected by saturation due to the so-called blooming effect. This effect has very low impact in the UV region (HCHO fitting interval range). The pixels affected by blooming effect have been flagged since version 02.00.00 of the Level 1b processor, operational since July 2021.

Metadata values exchanged (solved)

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 since the switch to version 02.00.00 of the Level 1b processor on July 2021.

Variable surface temperature with fill-value (solved in version 02.04.01)

Since version 02.04.01 the variable <code>surface_temperature</code> is filled in correctly (it was set to fill-value in the previous version).

4.3 Data Features

This section describes some characteristics of the data that might seem anomalous, however they are physically correct and not related to any problem.

Pixel geolocation around the North Pole (feature)

The solar irradiance is measured on a daily basis over the North Pole at a reference azimuth angle to remove seasonal effects on the measurements. To this end, a yaw manoeuvre is executed when the instrument is still in radiance mode, causing possible distortion on the scanlines observed during this manoeuvre (i.e. crossing scanlines, "bow-tie" ground pixel shape instead of rectangular). This occurs at most during the last 26 seconds of radiance measurements in few orbits (7-9 per week). Though this may seem anomalous, it is physically correct, and not related to any problem on the data geolocation.

4.4 Mission Operations Change

A change in the Copernicus Sentinel-5P operations scenario, increasing the spatial resolution from 7.0 km to 5.5 km along track for all measurements, became operational starting from 6 August 2019, orbit 9388.

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 have been no data format changes in the current release of the processor.

7 Product Availability

All S5P/TROPOMI data are available on the Copernicus Data Space Ecosystem https://dataspace.copernicus.eu

The list of major mission data gaps due to acquisition faults or satellite/instrument disruption is available at <u>https://sentinel.esa.int/web/sentinel/missions/sentinel-5p/mission-status</u>. For those periods the data are permanently lost.

RPRO dataset gaps: additional gaps are present on the reprocessed dataset (see Table 3) due to the unavailability of Level 0 (L0) input data during the full mission reprocessing campaign.

Orbit	Gap start time	Gap stop time
3546	20/06/2018 08:31:35	20/06/2018 08:51:37
9755	31/08/2019 23:46:24	01/09/2019 00:06:25
10782	12/11/2019 08:39:57	12/11/2019 09:29:55
19782	07/08/2021 18:09:52	07/08/2021 18:27:54
19785	07/08/2021 22:34:02	07/08/2021 22:50:52
20254	10/09/2021 00:01:42	10/09/2021 00:21:43

Table 3: Gaps on RPRO dataset due to the unavailability of L0 input data during the full mission reprocessing campaign

Information on data handling tools is available from the web page http://www.tropomi.eu/tools.

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

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 : <u>https://sentinel.esa.int/documents/247904/2474724/Sentinel-5P-Calibration-and-Validation-Plan.pdf</u>
[RD02]	Sentinel-5 precursor/TROPOMI Level 2 Algorithm Theoretical Basis Document Formaldehyde source: BIRA; ref: S5P- BIRA-L2- ATBD-400F; url: <u>https://sentinel.esa.int/documents/247904/2476257/Sentinel-5P-ATBD-HCHO-TROPOMI</u>
[RD03]	Sentinel-5 precursor/TROPOMI Level 2 Product User Manual Formaldehyde HCHO source: DLR; ref: S5P-L2-DLR-PUM-400F; url: <u>https://sentinel.esa.int/documents/247904/2474726/Sentinel-5P-Level-2-Product-User-Manual-Formaldehyde</u>
[RD04]	Algorithm theoretical basis document for the TROPOMI L01b data processor source : KNMI; ref: S5P-KNMI-L01B-0009-SD; url : <u>https://sentinels.copernicus.eu/documents/247904/2476257/Sentinel-5P-TROPOMI-Level-1B-ATBD</u>

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 <u>http://www.tropomi.eu/data-products</u>.

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
OCRA	Optical Cloud Recognition Algorithm
OMI	Ozone Monitoring Instrument
OMPS	Ozone Mapper and Profiling Suite
PRF	Product Readme File
PUM	Product User Manual
ROCVR	Routine Operations Consolidated Validation Report
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