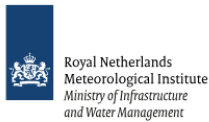




S5P Mission Performance Centre Formaldehyde [L2__HCHO__] Readme



| | | |
|-----------------|-------------------------------------------------------------------------|------------------------------------------------------------------------------|
| document number | S5P-MPC-BIRA-PRF-HCHO | |
| issue | 2.0 | |
| date | 2020-07-16 | |
| product version | V02.01.03 | |
| status | Released | |
| Prepared by | I. De Smedt (BIRA-IASB) Fabian Romahn (DLR) K.-U. Eichmann (IUPB) | MPC Product Lead MPC ESL-L2 processor lead MPC VAL Product Coordinator |
| Reviewed by | J.-C. Lambert (BIRA-IASB) D. Loyola (DLR) J. P. Veefkind (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 Compennolle (BIRA-IASB) Bavo Langerock (BIRA-IASB) K.L Chan (DLR) Lidia Saavedra de Miguel (ESA/Serco) | MPC ESL-VAL MPC ESL-VAL MPC ESL-L2 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 | MPC Product Lead / PRF Lead Editor | |
| | Angelika Dehn (ESA), 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 |

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____20190729T000912_20190729T001412_09273_01_020103_20200608T142937.nc

S5P_OFFL_L2_HCHO____20190729T081452_20190729T095819_09278_01_020103_20200520T132844.nc

The OFFL product has the following Digital Object Identifier (DOI): <https://doi.org/10.5270/S5P-vg1i7t0>

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.

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 version 1.x.x is compliant with the requirements as defined in the **S5P Calibration and Validation Plan** [RD01], see Table 1.

The quality of the processor version 2.x.x has been assessed on a dedicated Test Data Set.

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, and reports released after July 2020 include validation results based on processor version 2.x.x.

| Parameter | Data product | Vertical Resolution | Bias | Random |
|--------------|--------------|---------------------|------|--------------------------------------|
| Formaldehyde | HCHO | 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 |
|-------------------|----------------------------------------------------------------|----------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 02.01.03 | OFFL: orbit14239, 2020-07-13 NRTI: orbit 14285, 2020-07-16 | Current version | <ul style="list-style-type: none"> - 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 - Improve the parameter qa_value determination - Invalid values of geolocation_flags set to correct values - New variables added (see section Error! Reference source not found.) |
| 01.01.08 | OFFL: orbit 12432, 2020-03-07 NRTI: orbit 12482, 2020-03-11 | Orbit 14238, 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 | <ul style="list-style-type: none"> - 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 *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

Independent validation by the S5p MPC Cal/Val experts and the S5PVT concludes that the TROPOMI formaldehyde column data version 1.x.x 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 -44% for direct comparisons, and -25% 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 25 NDACC FTIR stations with high HCHO columns (-31% for HCHO > 8 Pmolec/cm₂), while a positive bias is observed over clean FTIR sites (+26% 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>. They are issued quarterly, and reports released after July 2020 include validation results based on processor version 2.x.x.

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.

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`.

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.

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. Note that, after this operations change, the metadata/Attributes fields related to the spatial resolution, remain **unchanged** (hence not aligned to the improved resolution).

QA values (solved in version 02.01.03)

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

Sun glint (solved in version 02.01.03)

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. 2018, 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.

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

6.1.1 New Variables

```
/PRODUCT/SUPPORT_DATA/INPUT_DATA/northward_wind  
/PRODUCT/SUPPORT_DATA/INPUT_DATA/eastward_wind  
/PRODUCT/SUPPORT_DATA/INPUT_DATA/sea_ice_cover  
/PRODUCT/SUPPORT_DATA/INPUT_DATA/snow_cover  
/PRODUCT/SUPPORT_DATA/INPUT_DATA/tm5_tropopause_layer_index
```

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 |
| 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 |