S5P Glyoxal Product User Manual

Document number	S5P-BIRA-PUM-CHOCHO	
issue	1.1	
date	16/11/2023	
product version	01.00.01	
status	release	
Prepared by	J. Vlietinck (BIRA-IASB) C. Lerot (BIRA-IASB) T. Danckaert (BIRA-IASB)	CHOCHO product lead
Reviewed by		
Approved by		

Version History

Revision	Date	Author(s)	Description
1.0	3/10/2022	Jonas Vlietinck, Christophe Lerot	First release.
1.1	16/11/2023	Jonas Vlietinck, Thomas Danckaert	satellite geolocation variables added.

Acronyms

AMF Air Mass Factor.
BIRA-IASB Royal Belgian Institute for Space Aeronomy.
CF Climate and Forecast.
CHOCHO gloyxal.
DOAS Differential Optical Absorption Spectroscopy.
ECMWF European Centre for Medium-Range Weather Forecasts.
ESA European Space Agency.
GLYRETRO GLYoxal Retrievals from TROPOMI.
GOME2 Global Ozone Monitoring Experiment-2.
NetCDF Network Common Data Form.
NMVOC Non-Methane Volatile Organic Compounds.
OFFL Offline.
OMI Ozone Monitoring Instrument.
PAL Product Algorithm Laboratory.
PFS Product Format Specification.
PRF Product Readme File.
RS Reference Sector.
S5P Sentinel-5 Precursor.
SCD Slant Column Density.
TROPOMI Tropospheric Monitoring Instrument.
UUID Universal Unique Identifier.
VCD Vertical Column Density.

Contents

1 Introduction to the TROPOMI/S5p Glyoxal product.

Glyoxal is mostly produced in the atmosphere as an intermediate product in the oxidation of other NMVOC's. It is also directly emitted from fire events and combustion processes. Several attempts to establish a glyoxal global budget using chemistry transport models[1, 2, 3, 4, 5] have estimated the production from natural sources and fires to about 70% and from human activities to about 30%. Owing to its short lifetime (a few hours), elevated glyoxal concentrations are observed near emission sources, see fig. 1. Satellite observations of glyoxal contribute therefore to provide information on NMVOC emissions in support to air quality and chemistry-climate related studies. In addition, glyoxal is also known to significantly contribute to the total budget of secondary organic aerosols [6], which impact both air quality and climate forcing.

Glyoxal has three absorption bands in the visible spectral range that have been exploited to remotely retrieve information on its atmospheric abundance using the Differential Optical Absorption Spectroscopy method [7] applied to ground-based [8, 9, 10, 11], air-borne [12, 13], ship-borne [14, 15] and space-based instruments. The first global glyoxal tropospheric column observations from space have been realized by [16] using nadir measurements from the SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CartograpHY) instrument. Based on this pioneering work, different glyoxal data products were derived from GOME-2 [17, 18] and from OMI [19, 20]. All those different products rely on a similar DOAS approach, but generally differ from each other by the choice of the fit settings and of the auxiliary input data.

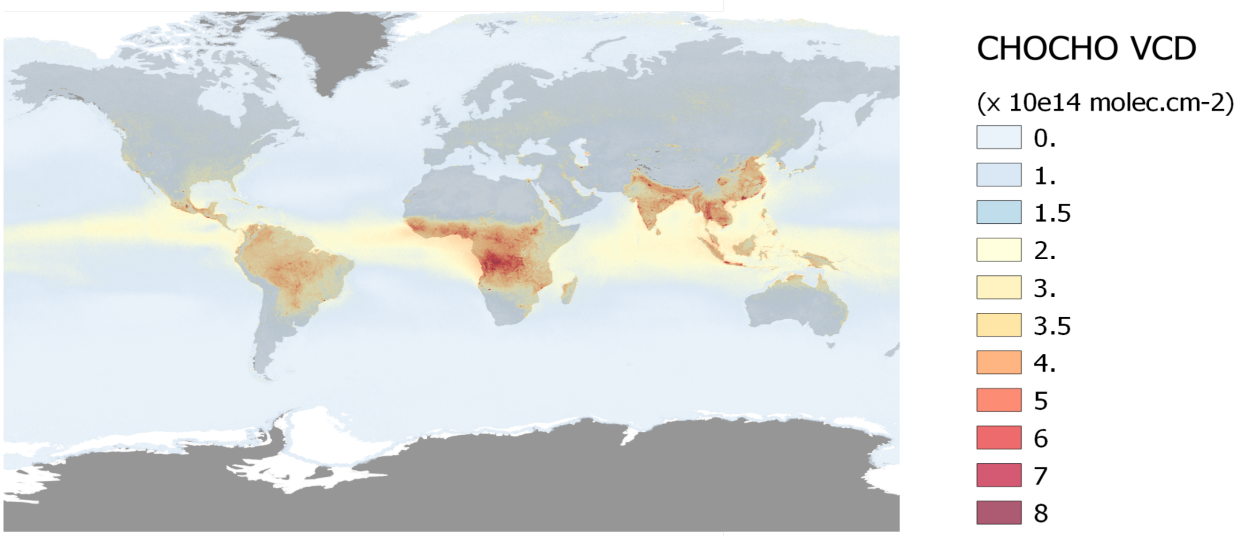

Figure 1: Mean glyoxal tropospheric vertical columns retrieved from three years of TROPOMI observations (2018-2020).

As part of the S5p+Innovation program, the BIRA-IASB scientific glyoxal retrieval algorithm has been further developed and applied to the TROPOMI L1 data. The product has been extensively compared with other satellite data products as well as with a few MAX-DOAS glyoxal data sets from stations in Asia and Europe. A detailed description of the algorithm and of the validation results can be found in the GLYRETRO ATBD [21] and Validation Report [22] as well as in [23]. A description of the GLYRETRO L2 and L3 file format can be found in [24]. The scientific requirements currently defined for glyoxal retrievals have been discussed in the requirement baseline document [25] and the satellite inter-comparison and validation exercises have shown that those requirements are generally met (for clear-sky scenes).

2 Introduction to the CHOCHO product files.

This document serves as a guide to understand the layout and format specifications of the CHOCHO product. For the construction of one L2__CHOCHO orbit file, a corresponding AUX_BGCHO_ and AUX_RARBD4 files are needed. This implies that there are three product types generated by the processor:

- **L2__CHOCHO** : For each orbit there is a L2 product, that contains the main results such as VCD and SCD.
- **AUX_BGCHO_** : The auxiliary product that contains SCDs averaged along the latitude and/or row dimensions within the RS.
- **AUX_RARBD4** : The auxiliary product that contains a radiance spectra averaged along the row dimension within the RS.

For an interpretation and explanation of the data in the L2 CHOCHO product, see [21]. An explanation of the filename structure of the product is given in section 3. The product is stored in a NetCDF4 binary format, following the CF-convention. Furthermore, the file format should be compliant with the guidelines provided in [26]. A description of the global attributes is provided in section 4. A full list of all variables that can be found in the L2 CHOCHO together with all the metadata is provided in section 5. Some advice on how to use the L2 product is given in section 6.

3 filename construction

For each orbit, a L2 file is generated, and has a filename described below. The production of L2 files rely on auxiliary files having their own filenames, also decribed below. The data in the auxiliary files are not limited to a specific orbit, but instead depends on a set of pixels from several L1 input files. Those pixels belong to a RS and are selected in two steps:

1. First, a time based selection is made to collect a set of subsequent L1 files (typically all L1 files from a given day).
2. The pixels from all those L1 files in step 1, are considered. The ones with latitude and longitude that falls into a RS are used to construct the aux. files.

3.1 L2__CHOCHO

S5P_<fileclass>_L2__CHOCHO_<start>_<end>_<orbit>_<coll>_<proc>_<mod>.nc

- **fileclass [4 characters** :] File class of the product. (example: PAL_)
- **start [YYYYMMDDThhmmss** :] start time of the orbit
- **end [YYYYMMDDThhmmss** :] end time of the orbit
- **orbit [5 digits** :] orbit number
- **coll [2 digits** :] collection id
- **proc [6 digits** :] processor version
- **mod [YYYYMMDDThhmmss** :] modification or creation time

3.2 AUX_BGCHO_

S5P_<fileclass>_AUX_BGCHO__<start>_<end>_<mod>.nc

- **fileclass [4 characters** :] File class of the product. (example: PAL_)
- **start [YYYYMMDDThhmmss** :] Corresponds to the earliest *start of the orbit*-time from the L1 files that are selected.
- **end [YYYYMMDDThhmmss** :] Corresponds to the latest *end of the orbit*-time from the L1 files that are selected.
- **mod [YYYYMMDDThhmmss** :] modification or creation time

3.3 AUX_RARBD4

S5P_<fileclass>_AUX_RARBD4_<start>_<end>_<mod>.nc

- **fileclass [4 characters** :] File class of the product. (example: PAL_)
- **start [YYYYMMDDThhmmss** :] Corresponds to the earliest *start of the orbit*-time from the L1 files that are selected.
- **end [YYYYMMDDThhmmss** :] Corresponds to the latest *end of the orbit*-time from the L1 files that are selected.
- **mod [YYYYMMDDThhmmss** :] modification or creation time

4 global attributes

In this section the global attributes in a product file are listed. The name of the attribute is provided together with the datatype. The static or dynamic nature of the attributes is also given. Static means that the attribute has the same values across all product files, dynamic means that the attributes values depends on the orbit of the product file.

4.1 L2__CHOCHO

Conventions [int32] (*static*)

CF-1.7 (Version of CF convtions that is followed.)

comments [string] (*dynamic*)

(Version of the python packages from which the processor is composed of).

- chocho-template : x.y.z
- chocho-amf : x.y.z
- chocho-bc : x.y.z
- chocho-rar : x.y.z

file_class [string] (*dynamic*)

File class of the product.

footprint [string] (*dynamic*)

GeoJSON format. Footprint of the product as a single GeoJSON string value.

history [string] (*dynamic*)

YYY-MM-DDThh:mm:ssZ chocho_main <Joborder filepath>, with the time string the generation date.

id [string] (*dynamic*)

Product name (filename without extension)

input_files [string] (*dynamic*)

List that contains the filenames of all inputs to the processor.

institution [string] (*static*)

BIRA-IASB

orbit [int32] (*dynamic*)

orbit number. (matches the orbit number in the filename)

processing_center [string] (*static*)

S5P-PAL

processor_version [string] (*dynamic*)

xx.yy.zz (version number of the processor)

source [string] (*static*)

Sentinel 5 precursor, TROPOMI, space-borne remote sensing, L2

summary [string] (*static*)

TROPOMI/S5P CHOCHO L2 data Swath 5.5x3.5km2

time_coverage_end [string] (*dynamic*)

YYYY-MM-DDThh:mm:ss.fffZ (Start time of last measurement in the product)

time_coverage_resolution [string] (*dynamic*)

PT<duration>S (duration in seconds of the scanline)

time_coverage_start [string] (*dynamic*)

YYYY-MM-DDThh:mm:ss.fffZ (Start time of first measurement in the product)

time_reference [string] (*static*)

YYYY-MM-DDThh:mm:ss.fffZ (Start of the day of the sensing time)

tracking_id [string] (*dynamic*)

UUID

4.2 AUX_BGCHO_

Conventions [int32] (*static*)

CF-1.7 (Version of CF convtions that is followed.)

comments [string] (*dynamic*)

(Version of the python packages from which the processor is composed of).

- chocho-template : x.y.z
- chocho-amf : x.y.z
- chocho-bc : x.y.z
- chocho-rar : x.y.z

file_class [string] (*dynamic*)

File class of the product.

footprint [string] (*dynamic*)

GeoJSON format. Footprint of the product as a single GeoJSON string value.

history [string] (*dynamic*)

YYY-MM-DDThh:mm:ssZ chocho_pre <Joborder filepath>, with the time string the generation date.

id [string] (*dynamic*)

Product name (filename without extension)

input_files [string] (*dynamic*)

List that contains the filenames of all inputs to the processor that were used in the construction of this product.

institution [string] (*static*)

BIRA-IASB

lat_bound [int64] (*dynamic*)

[lat_min,lat_max] Latitude boundaries of the reference sector.

lon_bound [int64] (*dynamic*)

[lon_min,lon_max] Longitude boundaries of the reference sector.

processing_center [string] (*static*)

S5P-PAL

processor_version [string] (*dynamic*)

xx.yy.zz (version number of the processor)

source [string] (*static*)

Sentinel 5 precursor, TROPOMI, space-borne remote sensing, L2

summary [string] (*static*)

TROPOMI/S5P CHOCHO L2 Background correction parameters

time_coverage_end [string] (*dynamic*)

YYYY-MM-DDThh:mm:ss.fffZ (Corresponds with the *end* time in the filename, see section 3.2)

time_coverage_resolution [string] (*dynamic*)

PT<duration>S (duration in seconds of the scanline)

time_coverage_start [string] (*dynamic*)

YYYY-MM-DDThh:mm:ss.fffZ (Corresponds with the *start* time in the filename, see section 3.2)

time_reference [string] (*static*)

YYYY-MM-DDThh:mm:ss.fffZ (Start of the day of the *time_coverage_start* attribute)

tracking_id [string] (*dynamic*)

UUID

4.3 AUX_RARBD4

Conventions [int32] (*static*)

CF-1.7 (Version of CF conventions that is followed.)

comments [string] (*dynamic*)

(Version of the python packages from which the processor is composed of).

- chocho-rar : x.y.z

file_class [string] (*dynamic*)

File class of the product.

history [string] (*dynamic*)

YYY-MM-DDThh:mm:ssZ chocho_pre <Joborder filepath>, with the time string the generation date.

id [string] (*dynamic*)

Product name (filename without extension)

input_files [string] (*dynamic*)

List that contains the filenames of all inputs to the processor that were needed to construct this product.

institution [string] (*static*)

BIRA-IASB

lat_bound [int64] (*dynamic*)

[lat_min,lat_max] Latitude boundaries of the reference sector.

lon_bound [int64] (*dynamic*)

[lon_min,lon_max] Longitude boundaries of the reference sector.

measurement date [string] (*dynamic*)

This attribute is the date (*YYYY/MM/DD*) of the day of the *start* time in the filename, see section 3.3

processing_center [string] (*static*)

S5P-PAL

processor_version [string] (*dynamic*)

xx.yy.zz (version number of the processor)

source [string] (*static*)

Radiance as reference file in APEX format for QDOAS based on daily averaged radiances

summary [string] (*static*)

TROPOMI/S5P CHOCHO L2 data Swath 5.5x3.5km2

time_coverage_end [string] (*dynamic*)

YYYY-MM-DDThh:mm:ss.fffZ (Corresponds with the *end* time in the filename, see section 3.3)

time_coverage_resolution [string] (*dynamic*)

PT<duration>S (duration in seconds of the scanline)

time_coverage_start [string] (*dynamic*)

YYYY-MM-DDThh:mm:ss.fffZ (Corresponds with the *start* time in the filename, see section 3.3)

time_reference [string] (*static*)

YYYY-MM-DDThh:mm:ss.fffZ (Start of the day of the *time_coverage_start* attribute)

tracking_id [string] (*dynamic*)

UUID

5 General structure of the L2 file content

5.1 L2__CHOCHO

5.2 /PRODUCT

corner [int32] (*corner*)

- **units :** 1
- **long_name :** pixel corner index
- **comment :** This coordinate variable defines the indices for the pixel corners; index starts a 0 (counter-clockwise, starting from south-western corner of the pixel in ascending part of the orbit).

delta_time [int32] (*time, scanline*)

- **long_name :** offset from reference start time of measurement
- **units :** milliseconds since 2020-08-30 00:00:00

glyoxal_tropospheric_vertical_column [float32] (*time, scanline, ground_pixel*)

- **units :** mol m-2
- **standard_name :** troposphere_mole_content_of_glyoxal
- **long_name :** vertical column of glyoxal
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude
- **multiplication_factor_to_convert_to_DU :** 2241.15
- **multiplication_factor_to_convert_to_molecules_percm2 :** 6.02214e+19

glyoxal_tropospheric_vertical_column_precision [float32] (*time, scanline, ground_pixel*)

- **units :** mol m-2
- **standard_name :** troposphere_mole_content_of glyoxal standard_error
- **long_name :** random error of vertical column density
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude
- **multiplication_factor_to_convert_to_DU :** 2241.15
- **multiplication_factor_to_convert_to_molecules_percm2 :** 6.02214e+19

ground_pixel [int32] (*ground_pixel*)

- **units :** 1
- **axis :** X
- **long_name :** across-track dimension index
- **comment :** This coordinate variable defines the indices across track, from west to east; index starts at 0

latitude [float32] (*time, scanline, ground_pixel*)

- **long_name :** pixel center latitude
- **units :** degrees_north
- **standard_name :** latitude
- **valid_min :** -90.0
- **valid_max :** 90.0
- **bounds :** /PRODUCT/SUPPORT_DATA/GEOLOCATIONS/latitude_bounds

layer [int32] (*layer*)

- **units :** 1
- **long_name :** layer dimension index

longitude [float32] (*time, scanline, ground_pixel*)

- **long_name :** pixel center longitude
- **units :** degrees_east
- **standard_name :** longitude
- **valid_min :** -180.0
- **valid_max :** 180.0
- **bounds :** /PRODUCT/SUPPORT_DATA/GEOLOCATIONS/longitude_bounds

qa_value [uint8] (*time, scanline, ground_pixel*)

- **units :** 1
- **scale_factor :** 0.01
- **add_offset :** 0.0
- **valid_min :** 0
- **valid_max :** 100
- **long_name :** data quality value
- **comment :** A continuous quality descriptor, varying between 0 (no data) and 1 (full quality data). Recommend to ignore data with qa_value < 0.5
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

scanline [int32] (*scanline*)

- **units :** 1
- **axis :** Y
- **long_name :** along-track dimension index
- **comment :** This coordinate variable defines the indices along track; index starts at 0

time [int32] (*time*)

- **units :** seconds since 2010-01-01 00:00:00
- **standard_name :** time
- **axis :** T
- **long_name :** reference time for the measurements
- **comment :** The time in this variable corresponds to the time in the time_reference global attribute

5.3 /PRODUCT/SUPPORT_DATA/GEOLOCATIONS

latitude_bounds [float32] (*time, scanline, ground_pixel, corner*)

- **units :** degrees_north

longitude_bounds [float32] (*time, scanline, ground_pixel, corner*)

- **units :** degrees_east

solar_azimuth_angle [float32] (*time, scanline, ground_pixel*)

- **long_name :** solar azimuth angle
- **standard_name :** solar_azimuth_angle
- **units :** degree
- **valid_min :** -180.0
- **valid_max :** 180.0
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude
- **comment :** Solar azimuth angle at the ground pixel location on the reference ellipsoid. Angle is measured clockwise from the North (East = 90, South = 180, West = 270)

solar_zenith_angle [float32] (*time, scanline, ground_pixel*)

- **long_name :** solar zenith angle
- **standard_name :** solar_zenith_angle
- **units :** degree
- **valid_min :** 0.0
- **valid_max :** 180.0
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude
- **comment :** Solar zenith angle at the ground pixel location on the reference ellipsoid. Angle is measured away from the vertical

viewing_azimuth_angle [float32] (*time, scanline, ground_pixel*)

- **long_name :** viewing azimuth angle
- **standard_name :** viewing_azimuth_angle
- **units :** degree
- **valid_min :** -180.0
- **valid_max :** 180.0
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude
- **comment :** Satellite azimuth angle at the ground pixel location on the reference ellipsoid. Angle is measured clockwise from the North (East = 90, South = 180, West = 270)

viewing_zenith_angle [float32] (*time, scanline, ground_pixel*)

- **long_name :** viewing zenith angle
- **standard_name :** viewing_zenith_angle
- **units :** degree
- **valid_min :** 0.0
- **valid_max :** 180.0
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude
- **comment :** Zenith angle of the satellite at the ground pixel location on the reference ellipsoid. Angle is measured away from the vertical

satellite_altitude [float32] (*time, scanline*)

- **long_name :** satellite altitude
- **units :** m
- **valid_min :** 700000.0
- **valid_max :** 900000.0
- **comment :** The altitude of the spacecraft relative to the WGS84 reference ellipsoid

satellite_latitude [float32] (*time, scanline*)

- **long_name :** sub-satellite latitude
- **units :** degrees_north
- **valid_min :** -90.0
- **valid_max :** 90.0
- **comment :** Latitude of the spacecraft sub-satellite point on the WGS84 reference ellipsoid

satellite_longitude [float32] (*time, scanline*)

- **long_name :** sub-satellite longitude
- **units :** degrees_east
- **valid_min :** -180.0
- **valid_max :** 180.0
- **comment :** Longitude of the spacecraft sub-satellite point on the WGS84 reference ellipsoid

satellite_orbit_phase [float32] (*time, scanline*)

- **long_name :** fractional satellite orbit phase
- **units :** 1
- **valid_min :** -0.02
- **valid_max :** 1.02
- **comment :** Relative offset (0.0 ... 1.0) of the measurement in the orbit

5.4 /PRODUCT/SUPPORT_DATA/INPUT_DATA

aerosol_index_354_388 [float32] (*time, scanline, ground_pixel*)

- **units :** 1
- **standard_name :** ultraviolet_aerosol_index
- **long_name :** Aerosol index from 388 and 354 nm
- **radiation_wavelength :** 354.0, 388.0
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

cloud_fraction_crb [float32] (*time, scanline, ground_pixel*)

- **units :** 1
- **long_name :** effective radiometric cloud fraction
- **source :** Extracted from S5p NO2 OFL product
- **comment :** Retrieved effective radiometric cloud fraction derived in NO2 fitting window
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

cloud_pressure_crb [float32] (*time, scanline, ground_pixel*)

- **units :** Pa
- **long_name :** cloud radiometric optical centroid pressure
- **source :** Extracted from S5p NO2 OFL product
- **comment :** Retrieved atmospheric pressure at the level of cloud.
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

land_ocean_flag [uint8] (*time, scanline, ground_pixel*)

- **units :** 1
- **long_name :** land/ocean mask
- **comment :** flag indicating whether center of ground pixel is over lands or oceans
- **source :** Derived from surface_classification variable in S5p Operational OFL NO2 Product
- **flag_meanings :** water land
- **flag_values :** 0UB, 1UB
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

snow_ice_flag [uint8] (*time, scanline, ground_pixel*)

- **units :** 1
- **long_name :** snow-ice mask
- **comment :** flag indicating snow/ice at center of ground pixel
- **source :** snow/ice variable in S5p Operational OFL NO2 product
- **flag_meanings :** snow-free_land sea_ice_1_percent sea_ice_2_percent sea_ice_3_percent sea_ice_4_percent sea_ice_5_percent sea_ice_6_percent sea_ice_7_percent sea_ice_8_percent sea_ice_9_percent sea_ice_10_percent sea_ice_11_percent sea_ice_12_percent sea_ice_13_percent sea_ice_14_percent sea_ice_15_percent sea_ice_16_percent sea_ice_17_percent sea_ice_18_percent sea_ice_19_percent sea_ice_20_percent sea_ice_21_percent sea_ice_22_percent sea_ice_23_percent sea_ice_24_percent sea_ice_25_percent sea_ice_26_percent sea_ice_27_percent sea_ice_28_percent sea_ice_29_percent sea_ice_30_percent sea_ice_31_percent sea_ice_32_percent sea_ice_33_percent sea_ice_34_percent sea_ice_35_percent sea_ice_36_percent sea_ice_37_percent sea_ice_38_percent sea_ice_39_percent sea_ice_40_percent sea_ice_41_percent sea_ice_42_percent sea_ice_43_percent sea_ice_44_percent sea_ice_45_percent sea_ice_46_percent sea_ice_47_percent sea_ice_48_percent sea_ice_49_percent sea_ice_50_percent sea_ice_51_percent sea_ice_52_percent sea_ice_53_percent sea_ice_54_percent sea_ice_55_percent sea_ice_56_percent sea_ice_57_percent sea_ice_58_percent sea_ice_59_percent sea_ice_60_percent sea_ice_61_percent sea_ice_62_percent sea_ice_63_percent sea_ice_64_percent sea_ice_65_percent sea_ice_66_percent sea_ice_67_percent sea_ice_68_percent sea_ice_69_percent sea_ice_70_percent sea_ice_71_percent sea_ice_72_percent sea_ice_73_percent sea_ice_74_percent sea_ice_75_percent sea_ice_76_percent sea_ice_77_percent sea_ice_78_percent sea_ice_79_percent sea_ice_80_percent sea_ice_81_percent sea_ice_82_percent sea_ice_83_percent sea_ice_84_percent sea_ice_85_percent sea_ice_86_percent sea_ice_87_percent sea_ice_88_percent sea_ice_89_percent sea_ice_90_percent sea_ice_91_percent sea_ice_92_percent sea_ice_93_percent sea_ice_94_percent sea_ice_95_percent sea_ice_96_percent sea_ice_97_percent sea_ice_98_percent sea_ice_99_percent sea_ice_100_percent permanent_ice snow mixed_pixels_at_coastlines suspect_ice_value corners ocean
- **flag_values :** 0UB, 1UB, 2UB, 3UB, 4UB, 5UB, 6UB, 7UB, 8UB, 9UB, 10UB, 11UB, 12UB, 13UB, 14UB, 15UB, 16UB, 17UB, 18UB, 19UB, 20UB, 21UB, 22UB, 23UB, 24UB, 25UB, 26UB, 27UB, 28UB, 29UB, 30UB, 31UB, 32UB, 33UB, 34UB, 35UB, 36UB, 37UB, 38UB, 39UB, 40UB, 41UB, 42UB, 43UB, 44UB, 45UB, 46UB, 47UB, 48UB, 49UB, 50UB, 51UB, 52UB, 53UB, 54UB, 55UB, 56UB, 57UB, 58UB, 59UB, 60UB, 61UB, 62UB, 63UB, 64UB, 65UB, 66UB, 67UB, 68UB, 69UB, 70UB, 71UB, 72UB, 73UB, 74UB, 75UB, 76UB, 77UB, 78UB, 79UB, 80UB, 81UB, 82UB, 83UB, 84UB, 85UB, 86UB, 87UB, 88UB, 89UB, 90UB, 91UB, 92UB, 93UB, 94UB, 95UB, 96UB, 97UB, 98UB, 99UB, 100UB, 101UB, 103UB, 252UB, 253UB, 254UB, 255UB
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

surface_albedo [float32] (*time, scanline, ground_pixel*)

- **units :** 1
- **long_name :** surface albedo
- **standard_name :** surface_albedo
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

surface_altitude [float32] (*time, scanline, ground_pixel*)

- **long_name :** surface altitude
- **standard_name :** surface_altitude
- **units :** m
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

surface_classification [uint8] (*time, scanline, ground_pixel*)

- **units :** 1
- **long_name :** Surface classification
- **source :** Surface classification variable extracted from S5p Operational OFL NO2 product
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

surface_pressure [float32] (*time, scanline, ground_pixel*)

- **units :** Pa
- **standard_name :** surface_air_pressure
- **long_name :** surface_air_pressure
- **source :**
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

5.5 /PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION

glyoxal_reference_sector_mean_air_mass_factor [float32] (*lat_nbins, ground_pixel*)

- **units :** 1

glyoxal_reference_sector_mean_air_mass_factor_trueness [float32] (*lat_nbins, ground_pixel*)

- **units :** 1

glyoxal_reference_sector_mean_model_scd [float32] (*lat_nbins, ground_pixel*)

- **units :** mol m-2

glyoxal_reference_sector_mean_scd [float32] (*lat_nbins, ground_pixel*)

- **units :** mol m-2

glyoxal_tropospheric_column_reference [float32] (*time*)

- **units :** mol m-2
- **long_name :** reference tropospheric column in the reference sector
- **standard name:** troposphere_mole_content_of_glyoxal
- **multiplication_factor_to_convert_to_DU :** 2241.15
- **multiplication_factor_to_convert_to_molecules_percm2 :** 6.02214e+19

glyoxal_tropospheric_column_reference_trueness [float32] (*time*)

- **units :** mol m-2
- **long_name :** Systematic error of the reference tropospheric column in the reference sector
- **multiplication_factor_to_convert_to_DU :** 2241.15
- **multiplication_factor_to_convert_to_molecules_percm2 :** 6.02214e+19

lat_nbins [float32] (*lat_nbins*)

number_of_reference_sector_mean_obs [int32] (*lat_nbins, ground_pixel*)

- **units :** 1

5.6 /PRODUCT/SUPPORT_DATA/DETAILED_RESULTS

averaging_kernel [float32] (*time, scanline, ground_pixel, layer*)

- **units :** 1
- **long_name :** total column averaging kernel
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

fitted_radiance_shift [float32] (*time, scanline, ground_pixel*)

- **units :** nm
- **long_name :** radiance wavelength shift from the doas fit
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

fitted_radiance_squeeze [float32] (*time, scanline, ground_pixel*)

- **units :** 1
- **long_name :** radiance wavelength squeeze/stretch from the doas fit
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

fitted_root_mean_square [float32] (*time, scanline, ground_pixel*)

- **units :** 1
- **long_name :** root mean square from the doas fit
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

fitted_slant_columns [float64] (*time, scanline, ground_pixel, number_of_slant_columns*)

- **units :** mol m-2
- **long_name :** slant columns of all absorbers: glyoxal, O3, NO2_220K, NO2_296K, H2O, O4, Ring, Liquid Water, Offset_0, Offset_1, resol, common residual, scene_heterogeneity1, scene_heterogeneity2
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude
- **multiplication_factor_to_convert_to_DU :** 2241.15
- **multiplication_factor_to_convert_to_molecules_percm2 :** 6.02214e+19

fitted_slant_columns_precision [float32] (*time, scanline, ground_pixel, number_of_slant_columns*)

- **units :** mol m-2
- **long_name :** slant columns errors of all absorbers: glyoxal, O3, NO2_220K, NO2_296K, H2O, O4, Ring, Liquid Water, Offset_0, Offset_1, resol, common residual, scene_heterogeneity1, scene_heterogeneity2
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude
- **multiplication_factor_to_convert_to_DU :** 2241.15
- **multiplication_factor_to_convert_to_molecules_percm2 :** 6.02214e+19

glyoxal_profile_apriori [float32] (*time, scanline, ground_pixel, layer*)

- **units :** 1
- **long_name :** a priori profile (vmr) interpolated in space and time on ground pixel
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

glyoxal_profile_apriori_pressure [float32] (*time, scanline, ground_pixel, layer*)

- **units :** Pa
- **long_name :** Pressure grid of a priori profile (Pa) interpolated in space and time on ground pixel
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

glyoxal_slant_column_corrected [float32] (*time, scanline, ground_pixel*)

- **units :** mol m-2
- **long_name :** corrected slant column density
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude
- **multiplication_factor_to_convert_to_DU :** 2241.15
- **multiplication_factor_to_convert_to_molecules_percm2 :** 6.02214e+19

glyoxal_slant_column_corrected_trueness [float32] (*time, scanline, ground_pixel*)

- **units :** mol m-2
- **long_name :** systematic error of the slant column density
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude
- **multiplication_factor_to_convert_to_DU :** 2241.15
- **multiplication_factor_to_convert_to_molecules_percm2 :** 6.02214e+19

glyoxal_tropospheric_air_mass_factor [float32] (*time, scanline, ground_pixel*)

- **units :** 1
- **long_name :** tropospheric air mass factor
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

glyoxal_tropospheric_air_mass_factor_kernel_trueness [float32] (*time, scanline, ground_pixel*)

- **units :** 1
- **long_name :** systematic error of the air mass factor for the kernel tropospheric glyoxal
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

glyoxal_tropospheric_air_mass_factor_precision [float32] (*time, scanline, ground_pixel*)

- **units :** 1
- **long_name :** random error of the tropospheric air mass factor
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

glyoxal_tropospheric_air_mass_factor_trueness [float32] (*time, scanline, ground_pixel*)

- **units :** 1
- **long_name :** systematic error of the tropospheric air mass factor
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

glyoxal_tropospheric_vertical_column_kernel_trueness [float32] (*time, scanline, ground_pixel*)

- **units :** mol m-2
- **long_name :** systematic error of the air mass factor for the kernel tropospheric glyoxal
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

glyoxal_tropospheric_vertical_column_trueness [float32] (*time, scanline, ground_pixel*)

- **units :** mol m-2
- **long_name :** Systematic error of vertical column density
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude
- **multiplication_factor_to_convert_to_DU :** 2241.15
- **multiplication_factor_to_convert_to_molecules_percm2 :** 6.02214e+19

number_of_slant_columns [int32] (*number_of_slant_columns*)

- **units :** 1
- **long_name :** number_of_slant_columns dimension index

scene_inhomogeneity_factor [float32] (*time, scanline, ground_pixel*)

- **units :** 1
- **long_name :** 0: homogeneous scene; large values: degree of heterogeneity (+/-1 max)
- **source :** Computed using small radiance measurements, given in L1 files.
- **coordinates :** /PRODUCT/longitude /PRODUCT/latitude

5.7 /PRODUCT/SUPPORT_DATA/DETAILED_RESULTS/ WAVELENGTH_CALIBRATIONS

calibration_subwindows_root_mean_square [float32] (*number_of_calibrations, number_of_subwindows*)

- **units :** 1
- **long_name :** calibration rms per subwindow

calibration_subwindows_shift [float32] (*number_of_calibrations, number_of_subwindows*)

- **units :** nm
- **long_name :** irradiance wavelengths shift fitted values per subwindow

calibration_subwindows_squeeze [float32] (*number_of_calibrations, number_of_subwindows*)

- **units :** 1
- **long_name :** irradiance wavelengths squeeze fitted values per subwindow

calibration_subwindows_wavelength [float32] (*number_of_calibrations, number_of_subwindows*)

- **units :** nm
- **long_name :** calibration wavelength center in each subwindow

number_of_calibrations [int32] (*number_of_calibrations*)

- **units :** 1
- **long_name :** number_of_calibrations dimension index

number_of_subwindows [int32] (*number_of_subwindows*)

- **units :** 1
- **long_name :** number_of_subwindows dimension index

5.8 AUX_BGCHO_

ground_pixel [int32] (*ground_pixel*)

Detector row dimension index.

glyoxal_reference_sector_mean_scd [float32] (*lat_nbins, ground_pixel*)

The averaged SCD over a RS as function of the latitude and row.

number_of_reference_sector_mean_obs [int32] (*lat_nbins, ground_pixel*)

Number of observations used in each latitude, ground_pixel bin for the averaging.

glyoxal_reference_sector_mean_air_mass_factor [float32] (*lat_nbins, ground_pixel*)

The averaged AMF over a RS as function of the latitude and row.

glyoxal_reference_sector_mean_model_scd [float32] (*lat_nbins, ground_pixel*)

The averaged SCD over a RS as function of the latitude and row.

lat_nbins [int32] (*lat_nbins*)

latitude dimension index.

glyoxal_reference_sector_mean_air_mass_factor_trueness [float32] (*lat_nbins, ground_pixel*)

The averaged AMF trueness over a RS as function of the latitude and row.

5.9 AUX_RARBD4

reference_wavelength [float64] (*col_dim, spectral_dim*)

The wavelength spectrum per detector row.

- **units :** 1e-09 m
- **standard_name :** radiation_wavelength

spectral_dim [int32] (*spectral_dim*)

Spectral dimension.

reference_radiance [float64] (*col_dim, spectral_dim*)

The averaged reference radiance spectrum per detector row.

- **units :** mol.m-2.nm-1.sr-1.s-1
- **long_name :** spectral photon radiance

col_dim [int32] (*col_dim*)

Detector row dimension index.

use_row [int32] (*col_dim*)

This variable is equal to one if a valid radiance has been obtained per row, or zero if not.

number_radiances [int32] (*col_dim*)

Number of radiances observations used for the averaging of the radiance for each detector row.

6 Recommendations for using the L2 CHOCHO product.

As mentioned before, the file format is netCDF-4, which is now a standard for Earth Observation missions. This format is versatile, flexible and permits the user to use netCDF-4 or HDF-5 APIs written in many data- analysis packages (e.g. IDL, MatLab, Python, C, C++,...) in order to read the data. This format also facilitates the visualization of the Geo-2D variables contained in the file with visualization tools such as Panoply http://www.giss.nasa.gov/tools/panoply/.

All variables contained in the L2 files are provided for all TROPOMI observations with a solar zenith angle less than 70°. No filter is applied by default for degraded fit quality, cloud or aerosol contamination, snow/ice cover... Instead, the user is recommended to use the qa_value variable to filter out the data. Keeping data with a qa_value equal or higher than 0.5 guarantees that only clear sky observations (cloud fraction less than or equal to 20%), not covered by snow/ice and with good fit quality are considered. This approach allows advanced users to implement their own filtering approach by making use of the different variables contained in the files. It is however reminded that no cloud correction is applied for the AMF computation. It is therefore essential to reject cloud-contaminated pixels, even if there is some flexibility on the cloud fraction thresholds to be used for this.

Advanced users may also exploit the column averaging kernels A and the a priori information provided in the L2 files when ingesting them in their own application. In DOAS, the column averaging kernel is computed as the altitude-dependent AMF (or Box-AMF) divided by the total AMF [27]. Averaging kernels can be used in the following ways:

- When comparing TROPOMI column data with independent profile data from other instruments or models, discrepancies due to the assumptions made in the TROPOMI retrievals can be minimized by recomputing a column N_v^{indep} by multiplying the partial column profile from the independent source with the averaging kernel as:

$$N_v^{indep} = \sum_i A_l x_{indep,l} \tag{1}$$

- Alternatively, TROPOMI AMFs can also be recomputed with any available a priori profile using the averaging kernels as

$$M' = M \frac{\sum_l A_l x'_l}{\sum_l x'_l} \tag{2}$$

  where M is the AMF provided in the L2 file based on the a priori profile used in the retrieval algorithm, and M' is the new AMF corresponding to the new a priori profile

where M is the AMF provided in the L2 file based on the a priori profile used in the retrieval algorithm, and M' is the new AMF corresponding to the new a priori profile.

The L2 product includes estimates for the random and systematic errors of the retrieved glyoxal tropospheric vertical column, namely the *glyoxal_tropospheric_vertical_column_precision* and the *glyoxal_tropospheric_vertical_column_trueness*. The first term is estimated by propagation of the instrumental noise through the algorithm. The second term results from all systematic errors in the different steps of the algorithm due to input data uncertainty, forward model and assumption errors .... Please refer to [21, 23] for a comprehensive description of those errors. Among the systematic errors, a significant component is the smoothing error, i.e. the error caused by the use of inexact a priori profile information in the AMF computation. As described above, this a priori profile can be replaced by any other external information by making use of the averaging kernels. If this new information can be considered as accurate, the smoothing error should become small. This motivates the provision of a total systematic error computed without this smoothing error (*glyoxal_tropospheric_vertical_column_kernel_trueness*).

The total error σ_{N_v} combining both the random and systematic errors can be computed for single observations as:

$$\sigma_{N_v} = \sqrt{\sigma^2_{N_v,rand} + \sigma^2_{N_v,syst}} \tag{3}$$

Glyoxal being a weak absorber, the random component is large and dominates for single observations. The product noise generally requires averaging of multiple observations in time and/or space to unambiguously detect real glyoxal signal. Only in case of extreme fire events, the measured glyoxal columns may occasionally exceed the detection limit for single observations. When combining N observations, the random component of the error is reduced and the total error may be computed as:

$$\sigma_{N_v} = \sqrt{\frac{\sigma^2_{N_v,rand}}{N} + \sigma^2_{N_v,syst}} \tag{4}$$

References

[1] H. Cao, T. M. Fu, L. Zhang, D. K. Henze, C. C. Miller, C. Lerot, G. G. Abad, I. De Smedt, Q. Zhang, M. Van Roozendael, F. Hendrick, K. Chance, J. Li, J. Zheng, and Y. Zha, "Adjoint inversion of chinese non-methane volatile organic compound emissions using space-based observations of formaldehyde and glyoxal," *Atmos. Chem. Phys*, vol. 18(20), pp. 15 017–15 046, 2018.

[2] T. Stavrakou, J.-F. Müller, I. De Smedt, M. Van Roozendael, M. Kanakidou, M. Vrekoussis, F. Wittrock, A. Richter, and J. P. Burrows, "The continental source of glyoxal estimated by the synergistic use of spaceborne measurements and inverse modelling," *Atmospheric Chemistry and Physics*, vol. 9, no. 21, pp. 8431–8446, 2009.

[3] S. Myriokefalitakis, M. Vrekoussis, K. Tsigaridis, F. Wittrock, A. Richter, C. Brühl, R. Volkamer, J. P. Burrows, and M. Kanakidou, "The influence of natural and anthropogenic secondary sources on the glyoxal global distribution," *Atmospheric Chemistry and Physics Discussions*, vol. 8, no. 1, pp. 1673–1708, Jan. 2008.

[4] T.-M. Fu, D. Jacob, F. Wittrock, J. Burrows, M. Vrekoussis, and D. Henze, "Global budgets of atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols," *Journal of Geophysical Research*, vol. 113, 08 2008.

[5] J. Li, J. Mao, K.-E. Min, R. A. Washenfelder, S. S. Brown, J. Kaiser, F. N. Keutsch, R. Volkamer, G. M. Wolfe, T. F. Hanisco, I. B. Pollack, T. B. Ryerson, M. Graus, J. B. Gilman, B. M. Lerner, C. Warneke, J. A. de Gouw, A. M. Middlebrook, J. Liao, A. Welti, B. H. Henderson, V. F. McNeill, S. R. Hall, K. Ullmann, L. J. Donner, F. Paulot, and L. W. Horowitz, "Observational constraints on glyoxal production from isoprene oxidation and its contribution to organic aerosol over the southeast united states," *Journal of Geophysical Research: Atmospheres*, vol. 121, no. 16, pp. 9849–9861, 2016.

[6] R. A. Washenfelder, C. J. Young, S. S. Brown, W. M. Angevine, E. L. Atlas, D. R. Blake, D. M. Bon, M. J. Cubison, J. A. de Gouw, S. Dusanter, J. Flynn, J. B. Gilman, M. Graus, S. Griffith, N. Grossberg, P. L. Hayes, J. L. Jimenez, W. C. Kuster, B. L. Lefer, I. B. Pollack, T. B. Ryerson, H. Stark, P. S. Stevens, and M. K. Trainer, "The glyoxal budget and its contribution to organic aerosol for los angeles, california, during calnex 2010," *J. Geophys. Res*, vol. 116(D21), p. D00V02, 2011.

[7] U. Platt and J. Stutz, *Differential Optical Absorption Spectroscopy: Principles and Applications*, ser. Physics of Earth and Space Environments. Springer Berlin Heidelberg, 2008.

[8] Z. Javed, C. Liu, M. F. Khokhar, W. Tan, H. Liu, C. Xing, X. Ji, A. Tanvir, Q. Hong, O. Sandhu *et al.*, "Ground-based max-doas observations of chocho and hcho in beijing and baoding, china," *Remote Sensing*, vol. 11, no. 13, p. 1524, 2019.

[9] N. Benavent, D. Garcia-Nieto, S. Wang, and A. Saiz-Lopez, "Max-doas measurements and vertical profiles of glyoxal and formaldehyde in madrid, spain," *Atmospheric Environment*, vol. 199, pp. 357–367, 2019.

[10] H. Hoque, H. Irie, and A. Damiani, "First max-doas observations of formaldehyde and glyoxal in phimai, thailand," *Journal of Geophysical Research: Atmospheres*, vol. 123, no. 17, pp. 9957–9975, 2018.

[11] S. F. Schreier, A. Richter, E. Peters, M. Ostendorf, A. W. Schmalwieser, P. Weihs, and J. P. Burrows, "Dual ground-based max-doas observations in vienna, austria: Evaluation of horizontal and temporal no2, hcho, and chocho distributions and comparison with independent data sets," *Atmospheric Environment: X*, vol. 5, p. 100059, 2020.

[12] R. Volkamer, S. Baidar, T. L. Campos, S. Coburn, J. P. DiGangi, B. Dix, E. W. Eloranta, T. K. Koenig, B. Morley, I. Ortega *et al.*, "Aircraft measurements of bro, io, glyoxal, no 2, h 2 o, o 2–o 2 and aerosol extinction profiles in the tropics: Comparison with aircraft-/ship-based in situ and lidar measurements," *Atmospheric Measurement Techniques*, vol. 8, no. 5, pp. 2121–2148, 2015.

[13] F. Kluge, T. Hüneke, M. Knecht, M. Lichtenstern, M. Rotermund, H. Schlager, B. Schreiner, and K. Pfeilsticker, "Profiling of formaldehyde, glyoxal, methylglyoxal, and co over the amazon: normalized excess mixing ratios and related emission factors in biomass burning plumes," *Atmospheric Chemistry and Physics*, vol. 20, no. 20, pp. 12 363–12 389, 2020.

[14] L. K. Behrens, A. Hilboll, A. Richter, E. Peters, L. Alvarado, A. B. Kalisz Hedegaard, F. Wittrock, J. P. Burrows, and M. Vrekoussis, "Detection of outflow of formaldehyde and glyoxal from the african continent to the atlantic ocean with a max-doas instrument," *Atmospheric Chemistry and Physics*, vol. 19, no. 15, pp. 10 257–10 278, 2019.

[15] R. Sinreich, S. Coburn, B. Dix, and R. Volkamer, "Ship-based detection of glyoxal over the remote tropical pacific ocean," *Atmospheric Chemistry and Physics*, vol. 10, no. 23, pp. 11 359–11 371, 2010.

[16] F. Wittrock, A. Richter, H. Oetjen, J. P. Burrows, M. Kanakidou, S. Myriokefalitakis, R. Volkamer, S. Beirle, U. Platt, and T. Wagner, "Simultaneous global observations of glyoxal and formaldehyde from space," *Geophysical Research Letters*, vol. 33, no. 16, 2006.

[17] C. Lerot, T. Stavrakou, I. De Smedt, J.-F. Müller, and M. Van Roozendael, "Glyoxal vertical columns from gome-2 backscattered light measurements and comparisons with a global model," *Atmospheric Chemistry and Physics*, vol. 10, no. 24, pp. 12 059–12 072, 2010.

[18] M. Vrekoussis, F. Wittrock, A. Richter, and J. Burrows, "Temporal and spatial variability of glyoxal as observed from space," *Atmospheric Chemistry and Physics*, vol. 9, no. 13, pp. 4485–4504, 2009.

[19] L. Alvarado, A. Richter, M. Vrekoussis, F. Wittrock, A. Hilboll, S. Schreier, and J. Burrows, "An improved glyoxal retrieval from omi measurements," *Atmospheric Measurement Techniques*, vol. 7, no. 12, pp. 4133–4150, 2014.

[20] C. Chan Miller, G. Gonzalez Abad, H. Wang, X. Liu, T. Kurosu, D. Jacob, and K. Chance, "Glyoxal retrieval from the ozone monitoring instrument," *Atmospheric Measurement Techniques*, vol. 7, no. 11, pp. 3891–3907, 2014.

[21] *Sentinel-5p+Innovation: Theme 1; Glyoxal Retrievals from TROPOMI (GLYRETRO): Algorithm Theoretical Baseline Document*, **Source**: BIRA-IASB, **issue**: 3.1, **ref**: S5p+I_CHOCHO_BIRA_ATBD, **date**: 2021-12-22.

[22] *Sentinel-5p+Innovation: Theme 1; Glyoxal Retrievals from TROPOMI (GLYRETRO): Product Validation Report*, **Source**: IUP-Bremen, **issue**: 2.1, **ref**: S5p+I_CHOCHO_BIRA_VR, **date**: 2021-12-22.

[23] C. Lerot, F. Hendrick, M. Van Roozendael, L. Alvarado, A. Richter, I. De Smedt, N. Theys, J. Vlietinck, H. Yu, J. Van Gent *et al.*, "Glyoxal tropospheric column retrievals from tropomi–multi-satellite intercomparison and ground-based validation," *Atmospheric Measurement Techniques*, vol. 14, no. 12, pp. 7775–7807, 2021.

[24] *Sentinel-5p+Innovation: Theme 1; Glyoxal Retrievals from TROPOMI (GLYRETRO): Product User Manual*, **source**: BIRA-IASB, **issue**: 1.1, **ref**: S5p+I_CHOCHO_BIRA_PUM , **date**: 2021-12-22.

[25] *Sentinel-5p+Innovation: Theme 1; Glyoxal Retrievals from TROPOMI (GLYRETRO): Product User Manual*, **source**: BIRA-IASB, **issue**: 2.0, **ref**: S5p+I_CHOCHO_BIRA_RB, **date**: 2019-11-01.

[26] *S5P-PAL: Sentinel 5P Product Algorithm Laboratory L2 Processor File Format Guidelines*, **source**: S&T; **ref**: ST-ESA-S5P_PAL-L2FFG-001; **issue**: 1.3; **date**: 2022-05-11.

[27] K. F. Boersma, H. J. Eskes, and E. J. Brinksma, "Error analysis for tropospheric no2 retrieval from space," *Journal of Geophysical Research: Atmospheres*, vol. 109, no. D4, 2004.