

FRM4DOAS Level-1 data

Caroline FAYT, caroline.fayt@aeronomie.be

Martina FRIEDRICH, martina.m.friedrich@aeronomie.be

François HENDRICK, Francois.Hendrick@aeronomie.be

Michel VAN ROOZENDAEL, Michel.VanRoozendael@aeronomie.be

BIRA-IASB

First version: 12 December 2017

Minor changes : 17 April 2018 (scan_index removal + addition of station altitude)

Major changes : 13th June 2018 – this document accounts for recommendations and comments from project partners made at the AR meeting (MPIC, 8 May 2018). The purpose of these new modifications was also to better match with satellite data file structure.

Group names in upper case

Creation of a group RADIANCE with two subgroups (OBSERVATIONS and GEODATA)

measurement group → RADIANCE/OBSERVATIONS group

metadata group → RADIANCE/OBSERVATIONS (any data except angles)

→ RADIANCE/GEODATA (solar and viewing angles)

ANCILLARY/METEOROLOGICAL_DATA divided in two groups :

TEMPERATURE_PRESSURE for temperature/pressure profiles

CLOUD_INFORMATION for cloud information

ANCILLARY group completed with SURFACE_DATA and AEROSOL_DATA sub-groups

Dimensions :

- num_rec dimension → number_of_records
- det_size dimension → detector_size
- additional dimensions for new data (cloud, aerosol)

Attributes for variables :

- description → long_name
- standard_name should be provided if known

Variables

- azimuth_viewing_angle → viewing_azimuth_angle
- elevation_viewing_angle → viewing_elevation_angle

Units :

- Latitude and longitude : resp. degree_north and degree_east instead of degree
- All wavelengths grids should always be given in “air”
- Radiance (sum of co-added spectra) should be in count number
- Time should always be specified in UT

Be careful to use variables without terminating s :

for example,

- for groups : RADIANCE, WAVELENGTH except OBSERVATIONS
- for variables : slit_function_relative_wavelength, slit_function_measured_wavelength

Minor changes: 04 November 2019 (time of ancillary data (meteo, cloud, aerosol) is now expressed in UT fractional day of year instead of UT fractional hour)

Minor changes: 30 January 2020 (adding the surface_temperature and surface_pressure variables in the ANCILLARY/METEOROLOGICAL_DATA/TEMPERATURE_PRESSURE group)

Minor changes: 21 January 2021 (the slit function dimensions in this document were not correct)

Minor changes: 7 April 2022

- Fill values of the ‘radiance_quality_flag’, ‘datetime_start’, ‘datetime_end’, and ‘measurement_type’ variables which were set to 1, 0, 0, 0, respectively, are now all set to -1.
- Update of the list of instruments

Miscellaneous

For 2D instruments with several azimuth observations measured the same time, it is better to separate in different files (use the channel number for the different azimuth and add a channel description in attributes). It should be the same for instruments that control a grating to alternate different wavelength ranges in

Introduction

Level-1 data shall consist in calibrated radiance spectra, dark current and offset corrected. The recommended format of these files is “netCDF”. There should be one file per measurement day. The proposed format offers the data submitters the possibility to include key data (e.g. the reference spectrum and the slit function) and ancillary data (e.g., pressure and temperature profiles).

This format will be used first by QDOAS for DOAS retrieval. The output data generated by QDOAS will be submitted to QA/QC procedure that will add quality flags to the data and will complete slant columns with the relevant ancillary information for profiling algorithms (temperature and pressure profiles if existing, geolocation and altitude of the instrument). The proposed file tries to follow the CF-1.6 convention.

Global attributes

- **Conventions** : the format convention (recommended to be CF compliant)
- **title** : Description of the data set (e.g., “Level-1 data for the FRM4DOAS MAXDOAS processor”)
- **source** : Description of the origin of the data
- **instrument_number** : unique number identifying the instrument (will be provided by BIRA)
- **instrument_channel** : channel number (depending on the instrument; to provide later in the metadata; for example, 1 for VIS, 2 for UV)
- **instrument_type** : maxdoas or zenith
- **institution** : the affiliation name (should be identical to the affiliation name in the file name in order to be able to reconstruct the L1 file name if necessary)
- **pi_name, pi_email** : the name and email address of the principal investigator
- **do_name, do_email** : the name and email address of the data originator
- **ds_name, ds_email** : the name and email address of the data submitter
- **station_name** : the name of the station in upper case
- **time_coverage_start** : starting date and time retrieved from the file name in the example below : 20180415T041746Z)
- **time_coverage_end** : starting date and time retrieved from the file name (in the example below : 20180415T190933Z)
- **measurement_funding_source** : optional, measurements could be funded by a specific organization
- **project_name** : should be FRM4DOAS for the moment
- **file_name_prefix** : should be ESA for the moment
- **file_type** : should be L1 for this file
- **file_version** : the file version (should be identical to the file version in the file name in order to be able to reconstruct the L1 file name if necessary)
- **campaign_name** : to keep empty for stations; to fulfil with the campaign name if the measurements have been operated during a campaign (for example, CINDI-2, TROLIX)

NB : for campaigns, both `station_name` and `campaign_name` attributes should be given. In this specific case, the station name is either the name of the location where the campaign took place or a location id provided by the organizers (for example : R5, R9, C12... for TROLIX)

Groups

The following groups and subgroups are proposed in the netCDF file :

- **ANCILLARY** : for ancillary data;
 - **METEOROLOGICAL_DATA** : temperature and pressures profiles
 - **SURFACE_DATA** : surface albedo
 - **AEROSOL_DATA** : data on aerosols

- **INSTRUMENT_LOCATION** : latitude, longitude and altitudes (station and instrument)
 - The difference of altitude between station and instrument should give the relative height of the instrument w.r.t the ground.
- **KEYDATA** :
 - **REFERENCE_SPECTRUM** : reference spectrum (not mandatory)
 - **SLIT_FUNCTION** : measured slit function(s)
- **RADIANCE** : for measurements
 - **GEODATA** : solar and viewing zenith and azimuth angles
 - **OBSERVATIONS**
 - radiances, radiances errors if known, radiances quality flags, wavelengths and all information about measurements
- **metadata** : currently limited to data present in ASCII files proposed for CINDI-2 reprocessing but could be extended to any data relevant for ground-based measurements (for example, scan index, temperature of the detector, ...)

Dimensions

The two most important dimensions are :

- **detector_size** : the size of the detector
- **number_of_records** : the number of records

They are completed with :

- The dimensions for TP profiles
 - **tp_level_size** : the number of altitude levels
 - **tp_time_size** : the number of profiles (depending on time)
- The dimensions of the slit functions matrix
 - **slit_dimx** : the number of relative wavelengths (grid on which the slit function is defined)
 - **slit_dimy** : the number of measured slit functions (the wavelengths have to be provided)
- **datetime_size** : for date and time variables, year, month, day, hours, minutes, seconds and milliseconds are provided in separate columns
- **dim1_size** : generic 1-length dimension (for example, the list of wavelengths at which the slit functions are provided).
- **cloud_size** : for cloud data

- aerosol_time_size, aerosol_wavelength_size : for aerosol data

Variables

See tables below.

Name of files

The following convention based on CCI and GEOMS file naming is proposed :

ESA-FRM4DOAS-L1-BIRA.IASB-UCCLE-1670-1-20180415T041746Z-20180415T190933Z-fv001.nc

Prefix for the project : ESA-FRM4DOAS

Processing level code : in this case, L1 for Level-1 data

Affiliation : name of the institute (see table of instruments below for available names)

The name of the station : see the table of instruments below

Instrument number : see the table of instruments below

Channel number : in this example, 1 (the description will be provided in a metadata file later)

Indicative date and time coverage

fvxxx : file version number xxx

nc : extension for netCDF files

List of instruments currently processed by the central processing system and with data files submission on NDACC/RD:

instrument_number	station_name	affiliation_name
1669	XIANGHE	BIRA.IASB
1670	UCCLE	BIRA.IASB
1671	HARESTUA	BIRA.IASB
1672	NY.ALESUND	IUP
1673	BREMEN	IUP
1674	ATHENS	IUP
1675	CABAUW	KNMI
1676	MAINZ	MPIC
1677	LAUDER	NIWA
1678	NEUMAYER	UHEIDELBERG
1679	HEIDELBERG	UHEIDELBERG
1681	LA.REUNION.MAIDO	BIRA.IASB
1683	THESSALONIKI	AUTH
1684	DEBILT	KNMI
1686	THESSALONIKI	AUTH
1688	DEBILT	KNMI
1689	UCCLE	BIRA.IASB

1690	CABAUW	KNMI
1693	CABAUW	KNMI
1694	IZANA	INTA

Group: INSTRUMENT_LOCATION

Name <i>standard_name</i>	Dim	Type	Unit	Fill value	long_name	Mandatory
altitude <i>altitude</i>	dim1_size	float32	m	NaN	Altitude of the instrument above sea level	Y
latitude <i>latitude</i>	dim1_size	float32	degree_north	NaN	Latitude of the instrument (positive north)	Y
longitude <i>longitude</i>	dim1_size	float32	degree_east	NaN	Longitude of the instrument (positive east)	Y
altitude_of_station <i>ground_level_altitude</i>	dim1_size	float32	m	NaN	Altitude of the station above sea level	Y

Group : ANCILLARY/METEOROLOGICAL_DATA/TEMPERATURE_PRESSURE (mandatory : N; default : temperature/pressure profiles)

Name	Dim	Type	Unit	Fill value	long_name
altitude_level	tp_level_size	float32	km	NaN	altitude levels above sea level. Not needed if surface pressure and temperature are reported.
meteo_time	tp_time_size	float32	day	NaN	UT time in fractional day of year
pressure	tp_level_size, tp_time_size	float32	hPa	NaN	Pressure profiles measured at different times and for different levels of altitudes
temperature	tp_level_size, tp_time_size	float32	K	NaN	Temperature profiles measured at different times and for different levels of altitudes
surface_pressure	tp_time_size	float32	hPa	NaN	Surface pressure measured at different times. Not

					needed if pressure profiles are reported.
surface_temperature	tp_time_size	float32	K	NaN	Surface temperature measured at different times. Not needed if temperature profiles are reported.

Group : ANCILLARY/METEOROLOGICAL_DATA/CLOUD_INFORMATION (mandatory : N; default : cloud coverage and height)

Name <i>standard_name</i>	Dim	Type	Unit	Fill value	long_name
cloud_time	cloud_size	float32	day	NaN	UT time in fractional day of year
cloud_coverage <i>cloud_area_fraction</i>	cloud_size	float32	percent	NaN	cloud coverage
cloud_height <i>cloud_base_altitude</i>	cloud_size	float32	km	NaN	cloud height

Group : ANCILLARY/AEROSOL_DATA (mandatory : N)

Name <i>standard_name</i>	Dim	Type	Unit	Fill value	long_name
aerosol_time	aerosol_time_size	float32	day	NaN	UT time in fractional day of year
aerosol_wavelength	dim1_size, aerosol_wavelength_size	float32	nm	NaN	wavelength (in air)
aerosol_optical_depth <i>atmosphere_optical_thickness_due_to_ambient_aerosol_particles</i>	aerosol_time_size, aerosol_wavelength_size	float32		NaN	aerosol optical depth
angstrom_exponent	aerosol_time_size,	float32		NaN	angström exponent

<i>angstrom_exponent_of_ambient_aerosol_in_air</i>					
asymmetry_factor	aerosol_time_size, aerosol_wavelength_size	float32		NaN	asymmetry factor
single_scattering_albedo <i>single_scattering_albedo_in_air_due_to_ambient_aerosol_particles</i>	aerosol_time_size, aerosol_wavelength_size	float32		NaN	single scattering albedo

Group : ANCILLARY/SURFACE_DATA (mandatory : N; default : NaN)

Name <i>standard_name</i>	Dim	Type	Unit	Fill value	long_name
surface_albedo <i>surface_albedo</i>	dim1_size	float32		NaN	Surface albedo

Group : KEYDATA/SLIT_FUNCTION (mandatory : N; default : characterized by QDOAS)

Name	Dim	Type	Unit	Fill value	long_name
slit_function_relative_wavelength	slit_dimx	float32	nm	NaN	relative wavelength grid
slit_function_measured_wavelength	dim1_size,slit_dimy	float32	nm	NaN	wavelengths at which the slit function has been measured
slit_function	slit_dimx,slit_dimy	float32		NaN	measured slit function(s)

Group : KEYDATA/REFERENCE_SPECTRUM (this field is not mandatory, by default : automatically selected by QDOAS)

Name	Dim	Type	Unit	Fill value	long_name
reference_wavelength	detector_size	float32	nm	NaN	reference wavelength grid (in air)
reference_spectrum	detector_size	float32		NaN	Reference spectrum

Group : RADIANCE/OBSERVATIONS (mandatory)

Name	Dim	Type	Unit	Fill value	long_name	Mandatory
wavelength	number_of_records, detector_size	float32	nm	NaN	wavelength grid (in air)	Y
radiance	number_of_records, detector_size	float32		NaN	sum of co-added spectra (count number)	Y
radiance_error	number_of_records, detector_size	float32		NaN	error on sum of co-added spectra (count number)	N
radiance_quality_flag	number_of_records, detector_size	int16		-1	pixel quality flags (0 for bad pixel or 1 for correct pixel but probably to refine later)	Y
exposure_time	number_of_records	float32	s	NaN	exposure time	Y
number_of_coadded_spectra	number_of_records	int16		-1	number of co-added spectra	Y
datetime	number_of_records, datetime_size	int16		-1	measurement date and time (UT YYYY,MM,DD,hh,mm,ss,ms)	Y
datetime_start	number_of_records, datetime_size	int16		-1	start date and time (UT YYYY,MM,DD,hh,mm,ss,ms)	N
datetime_end	number_of_records,	int16		-1	end date and time	N

	datetime_size				(UT YYYY,MM,DD,hh,mm,ss,ms)	
total_acquisition_time	number_of_records	float32	s	NaN	total acquisition time (the total time the detector is exposed to light to produce the spectrum)	N
total_measurement_time	number_of_records	float32	s	NaN	total measurement time (should be the time difference in seconds between UTC time start and UTC time stop)	N
measurement_type	number_of_records	int16		-1	measurement type : 0-invalid 1-offaxis, 2-direct sun, 3-zenith, 7-almucantar, 11-horizon, 12-direct moon	Y

Group : RADIANCE/GEODATA

Name	Dim	Type	Unit	Fill value	long_name	Mandatory
viewing_elevation_angle	number_of_records	float32	degree	NaN	viewing elevation angle	Y
viewing_azimuth_angle	number_of_records	float32	degree	NaN	viewing azimuth angle 0..360, measured towards the east, from north	Y
solar_zenith_angle	number_of_records	float32	degree	NaN	solar zenith angle	Y
solar_azimuth_angle	number_of_records	float32	degree	NaN	solar azimuth angle 0..360, measured towards the east, from north	Y
moon_zenith_angle	number_of_records	float32	Degree	NaN	moon zenith angle	N

moon_azimuth_angle	number_of_records	float32	degree	NaN	moon azimuth angle 0..360, measured towards the east, from north	N
--------------------	-------------------	---------	--------	-----	---	---