vegetation parameters cci

CCI Vegetation

D2.1 Algorithm Theoretical Basis Document
Cycle 2 – pre-processing

Else Swinnen, Sarah Gebruers, Kris Vanhoof, Dominique Jolivet, Simon Blessing

November 2024

Distribution list

Author(s)	:	Else Swinnen, Sarah Gebruers, Kris Vanhoof, Dominique Jolivet, Simon Blessing
Reviewer(s)	:	Christiaan Van der Tol
Approver(s)	:	Clement Albergel
Issuing authority	:	VITO

Change record

Release	Date	Pages	Description of change	Editor(s)/Reviewer(s)
V1.0	3/12/2024	all	First version	Christiaan van der Tol

Executive summary

CCI+ Vegetation Parameters is part of the ESA Climate Change Initiative. It aims at the identification, development and improvement of algorithms for the consistent retrieval of vegetation ECVs LAI and fAPAR from multi-platform and multi-mission satellite data and interact with the user community to match their requirements. The work plan includes three cycles, in which different data sources are combined, the algorithms' scientific and operational maturity is increased, and user feedback is incorporated.

This document summarizes the pre-processing steps of all satellite sensor data that are used as input for the retrieval of the second Climate Research Data package (CRDP-2). All processing steps from the L1B input data to the selection of data for the observation windows are described, and choices motivated. Chapter 2 provides an overview of these steps, the sensors that have been used for CRDP-2 (VIIRS, AVHRR, SPOT-VGT, PROBA-V and OLCI) and the time span of these missions. Chapter 3 describes the auxiliary data that have been selected for each of these sensors (and per platform/mission). Chapter 4 describes the process of tiling and gridding: The multi-sensor data have been projected on a common regular (1km) grid. For OLCI this process included spatial aggregation. Chapter 5 describes the atmospheric correction, which was similar for all sensors in order to ensure consistency, with only a small difference for OLCI, for which a single atmospheric profile was used. Chapter 6 describes the processing per sensor, providing further details on sensor-specific choices in the use of input data and flags. The appendices provide some background information on the atmospheric correction coefficients and the atmospheric profiles used.

Table of Contents

LIST OF ACRONYMS

Acronym	Definition
2D	2 dimensional
6S	Second Simulation of the Satellite Signal in the Solar Spectrum
AC	Atmospheric Correction
ADAS	Atmospheric Data Assimilation System
AERONET	Aerosol Robotic Network
AOT	Aerosol Optical Thickness
ATBD	Algorithm Theoretical Basis Document
AVHRR	Advanced Very High Resolution Radiometer
B0, B2, B3	Spectral bands of the SPOT/VEGETATION sensor, in the blue, red and near infrared, respectively.
BC	Black Carbon aerosols
C3S	Copernicus Climate Change
CAMS	Copernicus Atmosphere Monitoring Service
CCI	Climate Change Initiative
CGLOPS	Copernicus Global Land Service
CGLS	Copernicus Global Land Service
CRDP	Climate Research Data Package
CRS	Coordinate Reference System
CUDA	Compute Unified Device Architecture
DEM	Digital Elevation Model
DNB	Day/Night Band
DU	Dust aerosols
EC	European Commission
ECV	Essential Climate Variables
EOS	Earth Observation Satellites
EPS	EUMETSAT Polar System
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
fAPAR	Fraction of Absorbed Photosynthetically Active Radiation
GCOS	Global Climate Observing System
GEOS	Goddard Earth Observing System
GPL	GNU Public Licence
GPU	Graphics Processing Unit
KNN	Nearest-Neighbour
LAI	Leaf Area Index
LSA SAF	Land Surface Analysis Satellite Application Facility
MERRA-2	Modern-Era Retrospective analysis for Research and Applications, Version 2
MetOp	Meteorological-Operational satellites
MIR	Medium InfraRed
MODIS	MODerate Imaging Spectrometer
MODIS	Moderate Resolution Imaging Spectroradiometer
MVCM	MODIS-VIIRS Cloud Mask
NASA	National Aeronautics and Space Agency

NetCDF	Network Common Data Form
NIR	Near InfraRed
NOAA	National Oceanic and Atmospheric Administration
NRT	Near Real Time
NRT	Near Real Time
OC	Organic Carbon aerosols
OLCI	Ocean and Land Colour Instrument
PROBA-V	Project for Onboard Autonomy – Vegetation instrument
RSB	Reflective Solar Bands
RSS	Residual Sum of Square
SA	Surface Albedo
SAA	Sun Azimuth Angle
SeaWIFS	Sea-viewing Wide Field-of-view Sensor
SMAC	Simplified Model for Atmospheric Corrections
SMAC-GPU	SMAC model implemented with GPU
SNPP	Suomi National Polar-orbiting Partnership
SPOT	Satellite Probatoire d'Observation de la Terre
SS	Sea Salt aerosols
SU	SUlphate aerosols
SWIR	Short-Wave InfraRed
SZA	Sun Zenith Angle
TEB	Thermal Emissive Bands
TIR	Thermal Infrared
TOA	Top Of the Atmosphere
TOC	Top Of Canopy
TOC-r	Top Of Canopy Reflectance
UTC	Universal Time Coordinates
VAA	View Azimuth Angle
VEGETATION	The medium resolution sensor onboard SPOT4 and SPOT5
VGT	VEGETATION sensor
VIIRS	Visible Infrared Imaging Radiometer Suite
VIS	Visible
VITO	Vlaamse Instelling voor Technologisch Onderzoek (Flemish Institute for Technological Research), Belgium
VP	Vegetation Parameters
VR	Validation Report
VZA	View Zenith Angle
WGS	World Geodetic System
WGS	World Geodetic System

LIST OF FIGURES

LIST OF TABLES

1 Introduction

1.1 Scope of this document

This document describes the input data and the pre-processing algorithms applied to the input data to generate the surface reflectance data that are used in the retrieval of LAI and FAPAR. The description of the retrieval algorithm is described in [VP-CCI_D2.1_ATBD_V2.0].

1.2 Related documents

Internal documents

Reference ID	Document
ID1	Climate Change Initiative Extension (CCI+) Phase 2 New ECVs: Vegetation Parameters – EXPRO+ (ITT)
VP-CCI_D2.1_ATBD_V2.0	Algorithm Theoretical Basis Document of CRDP-2, ESA, CCI+ Vegetation Parameters

External documents

Reference ID	Document
ED1	C3S ATBD of Surface Albedo, multi-sensor, http://dast.data.compute.cci2.ecmwf.int/documents/satellite-albedo/D1.3.4-v2.0_ATBD_CDR_SA_MULTI_SENSOR_v2.0_PRODUCTS_v1.1.pdf
ED2	CGLS ATBD Sentinel-3 OLCI and SLSTR atmospheric correction , CGLOPS1_ATBD_S3-AC-V1_I1.30
ED3	Sentinel-3 OLCI Product Data Format Specification – OLCI Level 1 products https://sentinel.esa.int/documents/247904/1872756/Sentinel-3-OLCI-Product-Data-Format-Specification-OLCI-Level-1
ED4	CGLOPS1 – Preliminary Product User Manual – S3 Top of Canopy (TOC) reflectance 333 m https://land.copernicus.eu/global/sites/cgls.vito.be/files/products/CGLOPS1_ATBD_S3-AC-V1_I1.30.pdf

1.3 General definitions

Top of atmosphere (TOA) reflectance is the ratio of light reflected by a surface to the amount of solar radiation incident on that surface, measured at the top of the atmosphere. It is a unitless quantity. Since it is measured at the top of the atmosphere it includes atmospheric effects due to the optical properties of clouds, aerosols, etc. Because remote sensing provides directional reflectance, it is calculated as the ratio of π times the directional radiance and the solar irradiance.

Top of canopy (TOC) reflectance is defined as the fraction of solar radiation that is reflected by a surface, measured at the top of canopy. It is also dimensionless. The difference with TOA reflectance, is that the effects of the atmosphere are removed in TOC reflectances.

Projection is the mapping of the Earth's 3D surface onto a flat 2D surface. Near-polar satellites orbit the Earth across the poles. Due to the rotation of the Earth, the observational tracks of these satellites have S-shapes when projected onto a 2D surface. A regularly used projection is the plate carrée projection. This is an equirectangular projection in which latitude and longitude lines are represented as equally spaced horizontal and vertical lines, respectively. Every projection also requires a datum which is reference system that uses a mathematical model to fit the shape of the Earth onto an ellipsoid.

Atmospheric correction is the removal of absorption and scattering effects of the atmosphere from reflectance values. The light traveling from the Sun to the Earth's surface and from the surface to the image sensor passes the Earth's atmosphere twice. Many molecules in the atmosphere absorb and scatter this light at various wavelengths, thus effecting the reflectance detected by the sensor. By using an atmospheric correction method on TOA reflectances, these effects are eliminated and TOC reflectances are obtained.

2 General overview

Figure 1 sketches out the structure for algorithm development and successively increasing data complexity in the three development cycles of this project. Cycle 1 uses existing, atmospherically corrected, and gridded TOC reflectances from SPOT-VGT1/2 and PROBA-V as input. These TOC reflectance datasets are intermediate products of surface albedo production chains with validated output. The ATBDs for these chains are listed as external documents [ED1] for VGT and Proba-V 1 km gridded TOC reflectances.

In Cycle 2, data from Metop-A, -B and -C AVHRR, SNPP and NOAA-20 VIIRS, and Sentinel-3 OLCI are added to the input for the joint retrieval of LAI and fAPAR (see Figure 2). For Metop-AVHRR and VIIRS, the pre-processing is defined, implemented, and performed in the CCI+ VP project. The Sentinel-3 OLCI surface reflectance data from the Copernicus Global Land Service (CGLOPS) [ED2] are used after resampling to 1 km spatial resolution.

All surface reflectance datasets are projected onto the same grid with a spatial resolution of 1x1 km.

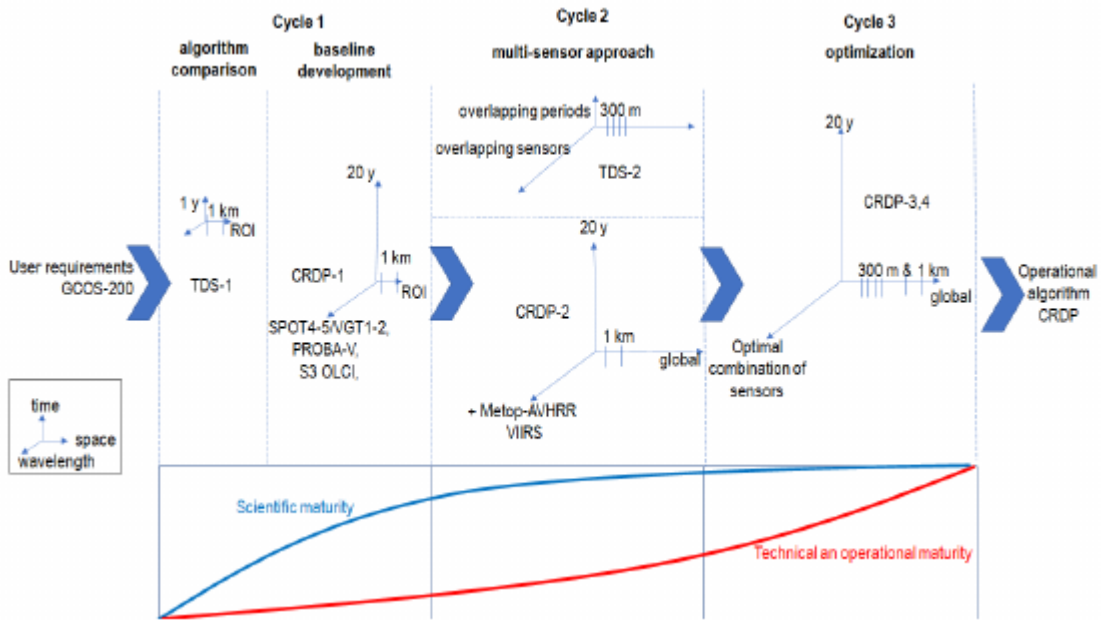

Figure 1: General concept of the three cycles, with progressive inclusion of sensors, spatial and temporal coverage and resolution, with the dimensions of the test datasets (TDS) and climate research data packages (CRDP)illustrated. The initially emphasis is on the implementation of an innovative approach, gradually shifting towards selection and optimization for an operational context

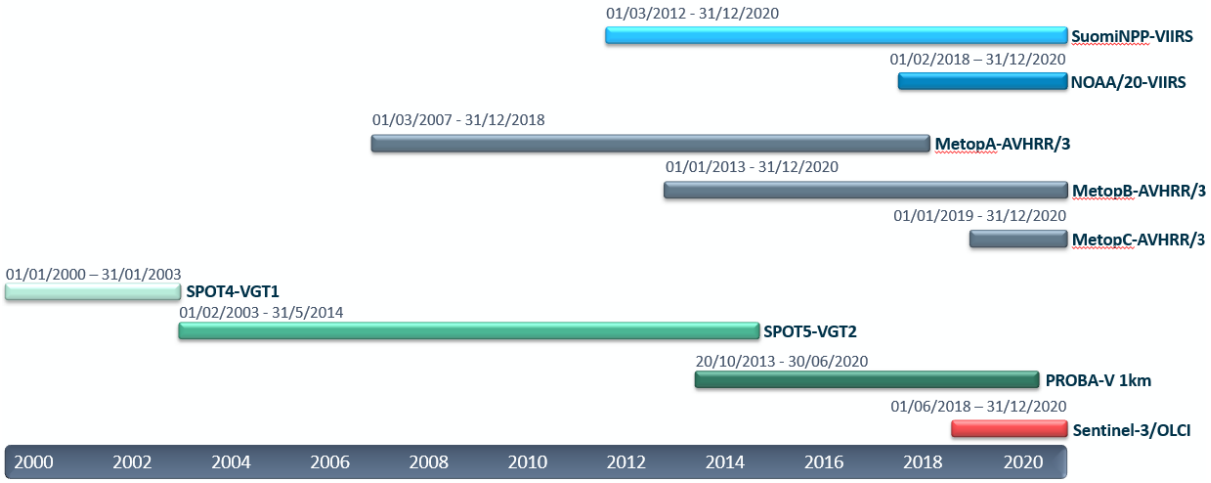

Figure 2: Available input data for CRDP-2.

Figure 3 shows the overall processing diagram of the CCI+ VP project. The red box indicates the pre-processing of the input datasets, which is the subject of this ATBD. The retrieval process starts with the selection from these input data. This is described in detail in [VP-CCI_D2.1_ATBD_V2.0].

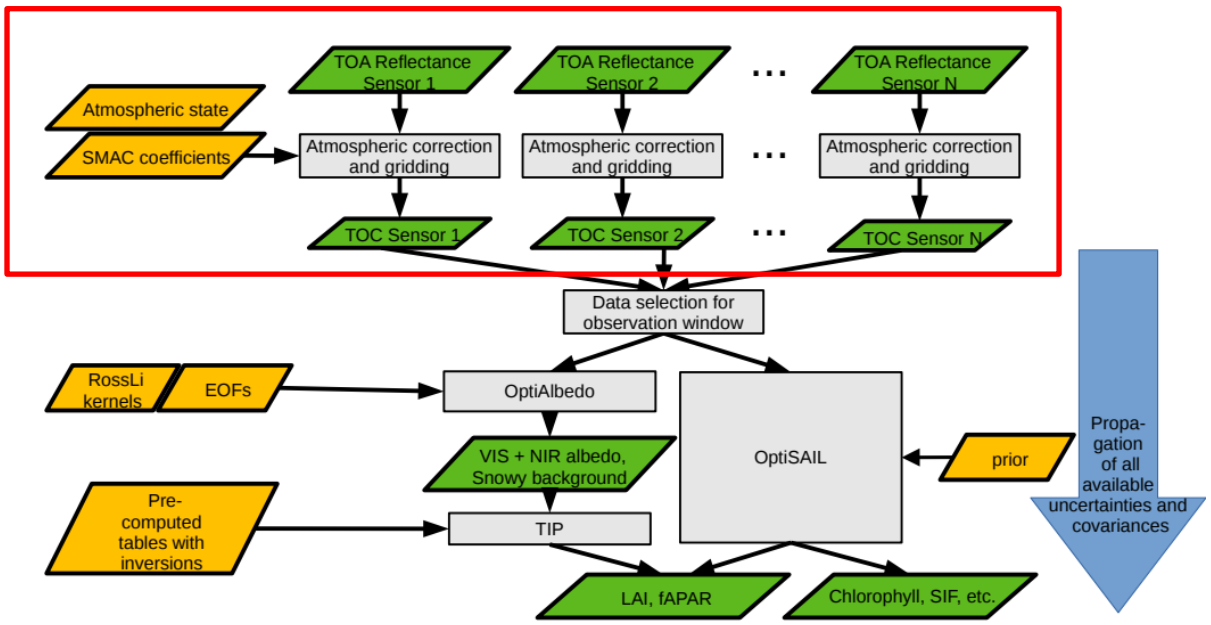

Figure 3: Processing diagram for CCI+ VP. The red rectangle indicates which part of the processing is summarized in this ATBD.

3 Input data

3.1 Sensors

The sensors for which the pre-processing steps are described in this ATBD are:

- AVHRR sensors on board Metop-A, B and C platforms for selected bands within the solar spectrum.

- VIIRS sensors on board JPSS-0, NOAA-20 and NOAA-21 for selected bands within the solar spectrum.

Data from other sensors that are used in the retrieval algorithm of LAI and FAPAR are:

- VGT1 and VGT2 sensors on board SPOT-4 and SPOT-5 respectively
- PROBA-V sensor onboard micro-satellite platform PROBA. PROBA-V has assured the succession of VGT sensors.
- OLCI sensors on board Sentinel 3 A and B, for selected bands within the solar spectrum.

The pre-processing of these sensors is described in [ED1] and [ED2]. Where applicable, reference to the processing approach is made in this document.

The period for which data from the sensors is available is represented in Figure 2.

3.1.1 Metop-AVHRR

The Meteorological-Operational (MetOp) satellites are operated by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT). It consists of three polar-orbiting satellites: Metop-A, Metop-B, and Metop-C launched in 2006, 2012, and 2018 respectively. Metop-A was decommissioned in 2021. The satellites follow a "morning orbit" with daytime registrations in the descending node. The local solar time of the overpasses is around 08:45 h for Metop-B and 09:30 h for Metop-A and -C, respectively. They provide detailed global observations of the atmosphere, the oceans and land.

One of the instruments on board the satellites is the Advanced Very High-Resolution Radiometer (AVHRR). It collects the Earth's outgoing radiation in five spectral bands, as shown in Table 1. Band 3 switches between shortwave infrared (SWIR) during daytime, band 3A, and medium infrared (MIR) observations during the night, band 3B. However, for vegetation monitoring only the daytime segments are used, thus band 3 always corresponds to the SWIR spectral range. The thermal infrared channels are also not used for vegetation parameters retrieval. The data have a spatial resolution of ±1 km at nadir and are used for global monitoring of land and sea surface temperature, cloud cover, snow and ice cover, soil moisture, and vegetation.

The raw observations are calibrated and transformed into TOA radiances by EUMETSAT, accurate "Lon/Lat-planes" are added with the geographical position of each pixel in the raw segment on the WGS84 geodetic datum, and a mask indicating the status of each observation (clear, cloud, snow) is added. The resulting data stream is cut into 3-minute segments (1080 scanlines).

Table 1: Spectral bands of Metop-AVHRR.

Band nr.	Bandwidth (µm)	Spectral domain	Band abbreviation
1	0.58 - 0.68	Shortwave	VIS (visual) or RED
2	0.725 - 1.00	Shortwave	NIR (near infrared)
3A	1.58 - 1.64	Shortwave	SWIR (shortwave infrared)
3B	3.55 - 3.93	Medium Infrared	MIR
4	10.3 - 11.3	Thermal Infrared	TIR_4
5	11.5 - 12.5	Thermal Infrared	TIR_5

The data from Metop-B AVHRR show a geolocation shift with respect to Metop-A and -C and the other satellite input data. The analysis of the geolocation shift in Toté et al. (2024) showed that the shift is variable in time and space and it is therefore not possible to correct for it easily.

3.1.2 VIIRS

The Visible Infrared Imaging Radiometer Suite (VIIRS) is one of five instruments onboard the Suomi National Polar-orbiting Partnership (NPP) satellite platform that was launched on October 28, 2011. VIIRS is also on board the National Oceanic and Atmospheric Administration (NOAA) weather satellite NOAA-20 that was launched in 2017. These two satellites are in the same polar orbit with NOAA-20 operating 50 minutes behind SNPP. The local solar time of the overpasses for VIIRS SNPP is 13:30 h. The VIIRS sensor was designed to extend and improve upon the series of measurements initiated by its predecessors, the Advanced Very High-Resolution Radiometer (AVHRR), the Moderate Resolution Imaging Spectroradiometer (MODIS), and the Sea-viewing Wide Field-of-view Sensor (SeaWIFS). VIIRS-derived data products are used to measure cloud and aerosol properties, ocean color, ocean and land surface temperature, ice movement and temperature, fires, and Earth's albedo.

The VIIRS instrument observes and collects global satellite observations that span the visible and infrared wavelengths across land, ocean, and atmosphere. A whiskbroom radiometer by design, it has 22 channels ranging from 0.41 µm to 12.01 µm. Five of these channels are high-resolution image bands or I-bands with a spatial resolution of 375 m at nadir. Of these five bands, three are Reflective Solar bands (RSB) and two are Thermal Emissive bands (TEB). Sixteen bands serve as moderate-resolution bands or M-bands with a resolution of 750 m across the entire scan. The M-bands comprise 11 RSB and five TEB bands. VIIRS also hosts a unique panchromatic Day/Night band (DNB), which is ultra-sensitive in low-light conditions that allows us to observe nighttime lights with better spatial and temporal resolutions compared to previously provided nighttime lights data by the Defense Meteorological Satellite Program. An overview is given in Table 2. In CCI only the 750 m M-bands are used.

The VIIRS Level-1B products contain calibrated radiance values, quality flags, uncertainty indices, and an array of related information that are contained in three data groups: observation data, scan line attributes, and diagnostics; data exist in separate products for each resolution (I-bands, M-bands, and DNB).

Table 2: Spectral bands of VIIRS (from VIIRS L1B User Guide[1]).

	Band	Wavelength (µm)	Spatial Res. at-Nadir (m)	Gain
	Image-resolution Bands			
RSBs	I1	0.600 – 0.680	375	Single
	I2	0.846 – 0.885	375	Single
	I3	1.580 – 1.640	375	Single
TEBs	I4	3.550 – 3.930	375	Single
	I5	10.50 – 12.40	375	Single
	Moderate-resolution Bands			
Reflective Solar Bands	M1	0.402 – 0.422	750	High/Low
	M2	0.436 – 0.454	750	High/Low
	M3	0.478 – 0.498	750	High/Low
	M4	0.545 – 0.565	750	High/Low
	M5	0.662 – 0.682	750	High/Low
	M6	0.739 – 0.754	750	Single
	M7	0.846 – 0.885	750	High/Low
	M8	1.230 – 1.250	750	Single
	M9	1.371 – 1.386	750	Single
	M10	1.580 – 1.640	750	Single
	M11	2.225 – 2.275	750	Single
Thermal Emissive Bands	M12	3.660 – 3.840	750	Single
	M13	3.973 – 4.128	750	High/Low
	M14	8.400 – 8.700	750	Single
	M15	10.263 – 11.263	750	Single
	M16	11.538 – 12.488	750	Single
	Day-Night band	0.500 – 0.900	750 [Across entire scan]	Low | Mid | High

3.2 Auxiliary data

3.2.1 Digital Elevation Model

The GTOPO30 (Global 30 Arc-Second Elevation) dataset is used for assigning a surface elevation to each pixel. It is downloaded as a set of binary files (https://lta.cr.usgs.gov/GTOPO30) and then re-assembled as a unique NetCDF file.

- Frequency: Static
- Spatial Grid: 2D, full extent of latitude from 90 degrees south to 90 degrees north, and the full extent of longitude from 180 degrees west to 180 degrees east. The horizontal grid spacing is 30-arc seconds.

1 https://ladsweb.modaps.eosdis.nasa.gov/api/v2/content/archives/Document%20Archive/Science%20Data%20Product%20Documentation/NASA_VIIRS_L1B_UG_August_2021.pdf

- Dimensions of 21,600 rows and 43,200 columns. The horizontal coordinate system is decimal degrees of latitude and longitude referenced to WGS84.
- The vertical units represent elevation above mean sea level (elev) and are expressed in meters. The values range from -407 to 8,752. It is complemented by the spatial standard deviation of the altitude (Delev in meters) computed in 3x3 pixel boxes.

3.2.2 Atmospheric Parameters

Depending on the sensors and the associated version of the SMAC algorithm two sources of ancillary data are used for atmospheric parameters.

- The Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) for the following sensors: VGT, PROBA-V, AVHRR and VIIRS
- Due to the near real time constraint of Copernicus Global Land Operations (CGLOPS) for S3-OLCI, atmospheric parameters come from CAMS NRT database, S3 input file auxiliary data and MERRA-2 monthly climatology.

3.2.2.1 VGT, PROBA-V, AVHRR and VIIRS sensors

The Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) is a NASA atmospheric reanalysis for the satellite era using the Goddard Earth Observing System Model, Version 5 (GEOS-5) with its Atmospheric Data Assimilation System (ADAS), version 5.12.4. It is the data source for all atmospheric parameters (https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/). The README file for the NetCDF MERRA-2 products is available here: (https://goldsmr4.gesdisc.eosdis.nasa.gov/data/MERRA2/M2I1NXINT.5.12.4/doc/MERRA2.README.pdf), while the full product specification is available here: (https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf).

The datasets are available from 1st January 1980 until present. They can be downloaded from several sites, for example (https://disc.sci.gsfc.nasa.gov/daac-bin/FTPSubset2.pl).

Two products are currently used for all atmospheric parameters:

- tavg1_2d_slv_Nx: 2-dimensional, 1-Hourly, Time-Averaged for physical parameters, ozone and water vapour.
- tavg1_2d_aer_Nx: 2-dimensional, 1-Hourly, Time-Averaged for aerosol diagnostics.

Both of them have the same spatial and temporal grids:

- Frequency: 24 1-hourly from 00:30 UTC (time averaged)
- Spatial Grid: 2D, single-level, global Lat-Lon regular grid
- Lat: 361 values from -90 to 90
- Lon: 576 values from -180 to 179.375

The atmospheric parameters used, and their sources, for the atmospheric corrections are listed in Table 3.

Table 3 : Auxiliary atmospheric parameters needed for the Atmospheric Corrections for AVHRR, VIIRS, PROBA-V and VGT.

Parameter	Symbol and unit	Source
Total Column Ozone	TO3 in Dobson units	MERRA2 tavg1_2d_slv_Nx products
Total Precipitable Water Vapor	TQV in $kg.m^{-2}$	
Sea-level Pressure	SLV in Pa	
Temperature above ground (10 m height)	T10M in Kelvin	

Aerosol Optical Thicknesses	TOTEXTTAU: the total aerosol extinction optical thickness at 550 nm. BCEXTTAU: the black carbon (BC) aerosol extinction optical thickness at 550 nm. OCEXTTAU: the organic carbon (OC) aerosol extinction optical thickness at 550 nm. DUEXTTAU: the dust (DU) aerosol extinction optical thickness at 550 nm. SUEXTTAU: the sulfate (SU) aerosol extinction optical thickness at 550 nm. SSEXTTAU: the sea salt (SS) aerosol extinction optical thickness at 550 nm	MERRA2 tavg1_2d_aer_Nx

3.2.2.2 Sentinel-3 OLCI

Table 4 lists the atmospheric parameters and their origins that are needed for the atmospheric correction of Sentinel-3 OLCI.

Table 4: Auxiliary atmospheric parameters needed for the Atmospheric Corrections for OLCI.

Parameter	Symbol and unit	Source
Total Column Ozone	TO3 in Dobson units	CAMS NRT
Total Precipitable Water Vapor	TQV in $kg.m^{-2}$	S3 input file (not available on CAMS NRT)
Sea-level Pressure	SLV in Pa	S3 input file (not available on CAMS NRT)
Temperature above ground (10 m height)	T10M in Kelvin	MERRA-2 monthly climatology
Aerosol Optical Thickness	TOTEXTTAU: Total Aerosol Optical Depth at 550 nm	CAMS NRT

The data of **Total column of ozone and the Aerosol Optical Thickness (AOT)** have been extracted from the CAMS NRT database. The Copernicus Atmosphere Monitoring Service (CAMS) is part of the Copernicus Program, which is an EU Program managed by the European Commission (EC) and implemented in partnership with the Member States, the European Space Agency (ESA), the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), the European Centre for Medium-Range Weather Forecasts (ECMWF), EU Agencies and Mercator Ocean. Near Real time products are needed for the operational processing within CGLOPS. These datasets are available from July 2012, well before the launch of Sentinel-3.

The product currently used for all atmospheric parameters is cams_nrt_surface_YYYYMMDD.grib, where YYYY is the year, MM is the month and DD is the day, retrieved through ftp download from: dissemination.ecmwf.int /DATA/CAMS_GLOBAL/<YYYYMMDD>00, with <YYYYMMDD> being the date of interest. The temporal resolution is 3 hours and the spatial resolution is about 40 km. Total column of water vapour, the sea-level surface pressure and surface temperature (T2m) from CAMS are provided with a delay of 5 days. Consequently, they cannot be used for NRT operations, but this is not a limitation for the Vegetation-CCI project, but it is relevant for possible future operational services..

The temporal resolution of CAMS NRT AOT and total column of ozone is 3 hours (00:00 UT, 03:00 UT, 06:00 UT, etc.). The CAMS NRT products are spatially and temporally interpolated to match with the Sentinel-3 overpass time.

Total column of water vapour and sea-level surface pressure come from S3-OLCI input files. These variables are given at pixel resolution.

The surface temperature comes from a climatology based upon a MERRA_2 monthly mean of 2018. It can also be downloaded at: https://disc.gsfc.nasa.gov/datasets?keywords=%22MERRA-2%22&page=1&source=Models%2FAnalyses%20MERRA-2

3.2.3 Selection of aerosol models

The last of the atmospheric parameters to be specified is the aerosol model. For OLCI the aerosol model is the continental one. For others sensors, the aerosol model is selected among 148 models (see Annex 2) by minimizing the distance in a 5-dimensional space corresponding to the 5 AOT ratios {Xu, u ∈ {DU, SU, OC, BC, SS }} of the 5 aerosols components Eq. 2.

$$X_u = \tau_{a;u}^{550} \Big/ \tau_a^{550} \qquad Eq.\ 1$$

The aerosol model number, iaero, is then calculated as follow

$$iaero = Min\left\{\sum_{u=1}^{5}\left(X_u - X_{u,i}\right)^2\right\}_{i=0,147} \qquad Eq.\ 2$$

The Xu,i basis for the 148 predefined models are stored in a text file and read once.

4 Projection and tiling

4.1 Projection and tiling overview

The projection step transforms a satellite image segment from its original viewing geometry into a target geometry that forms a regular grid in a defined coordinate reference system (CRS) with uniform spatial resolution for each pixel in each dimension. Tiling then splits the resulting image into tiles along predefined tiling grid lines.

In addition to radiance data and various other image layers, a source segment product contains a longitude and latitude image layer. Each pixel in the longitude/latitude images corresponds to a pixel in the other images, and thus labels each image coordinate (x,y) with a spatial coordinate (lon, lat). A coordinate (lon, lat) is assumed to refer to the center of an image pixel. Together the spatial coordinates typically form an irregular grid, with non-uniform distances between them.

The process is outlined in Figure 4. As a first step, the longitude and latitude pixel data are loaded into memory. If for any reason the data are stored in a down sampled form, it needs to be up sampled again to maintain pixel correspondence to the other image layers. Optionally, pixel data can be discarded at this point, for example when the view-zenith or solar-zenith angle exceeds a certain threshold.

Then a spatial index data structure is built from the segment longitude/latitude images. This spatial index enables a fast nearest-neighbour (KNN) search to find the pixel in the segment image data with the closest distance to a given (lon, lat) coordinate.

Next a source segment contour is derived from the segment's longitude/latitude images. Special care must be taken when images cross the -180° /180° longitude dateline, in which case the contour is split into parts.

The segment contour is then used to find all overlapping tiles in the tiling grid. From this point on, all processing is per tile and all tiles can be processed independently.

For each tile, a regular grid of spatial coordinates is generated, aligned with the tiling grid and taking into account the requested target resolution, these are the coordinates for the target image pixels:

$$lon = \{ x : x = ulx + res \cdot n,\ n \in \{0, 1, \ldots, W-1\} \}$$

$$lat = \{ y : y = uly - res \cdot n,\ n \in \{0, 1, \ldots, H-1\} \}$$

Where ulx is the longitude of the upper left corner of the tile, uly is the latitude of the upper left corner of the tile, res is the desired output resolution in degrees and $W \times H$ are the dimensions of the target tile image in number of pixels. Coordinates always refer to pixel centers.

Then for this grid of target coordinates, the spatial index is used to find the nearest source segment pixels. Target coordinates outside of the segment/tile contour are discarded before the search. Source pixels that have a distance farther than the source resolution from the target pixels are also discarded.

Finally, now that for each target pixel a source pixel has been identified, all image layers in the source segment product can be transformed and written into the target tile product through a series of indexed copying operations.

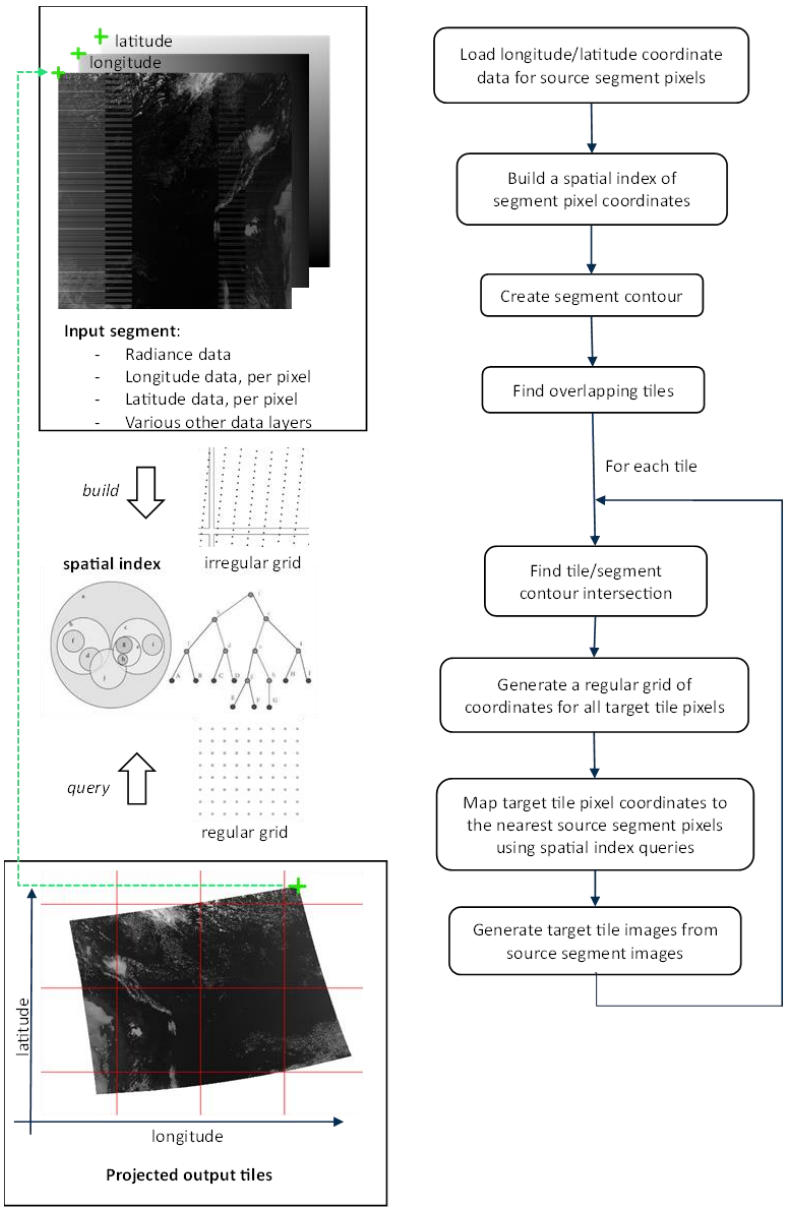

Figure 4: Outline of the projection and tiling processing.

4.2 Projection definition and method

The product is generated in a regular latitude/longitude grid (plate carrée projection) with the ellipsoïd WGS 1984 (Terrestrial radius=6378km). The resolution of the grid is 1/112°, which is about 1 km at the equator.

The reference is the centre of the pixel. It means that the longitude of the upper left corner of the pixel is (pixel_longitude – angular_resolution/2.). The products are provided in 10° x 10° tiles and per sites.

4.3 Tiling definition and method

The definition of the grid tiles for VIIRS and AVHRR is shown in Figure 5, and the one for Proba-V and VGT is displayed in Figure 6. Both tiling grids are the same in the longitudinal direction, but in the latitudinal direction, the Y00 tiles in the Proba-V/VGT grid correspond to the Y01 tiles in the VIIRS/AVHRR grid.

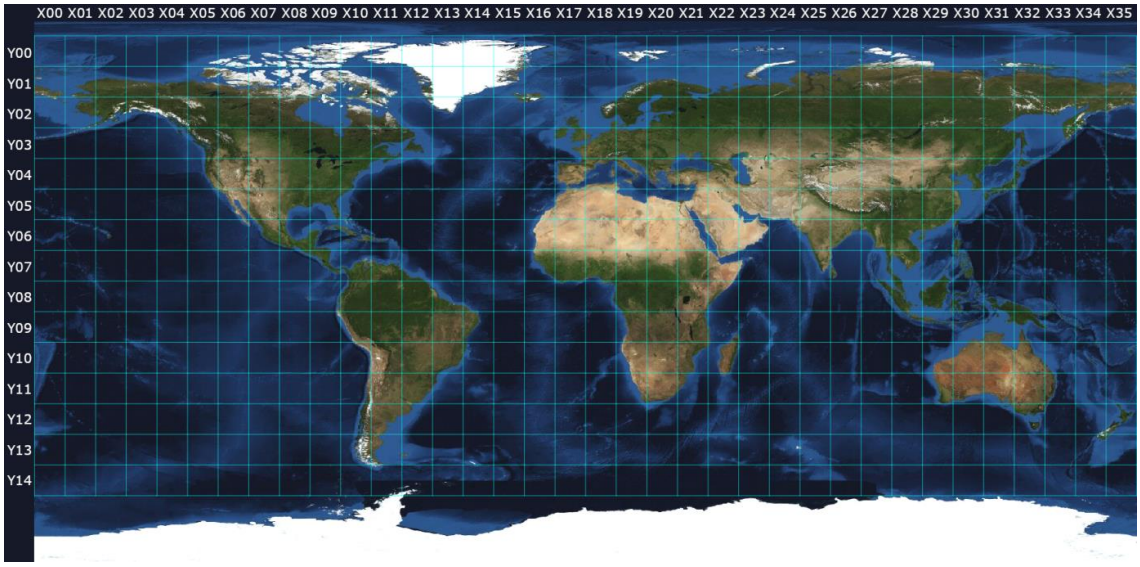

Figure 5: Grid tile definition for VIIRS and AVHRR.

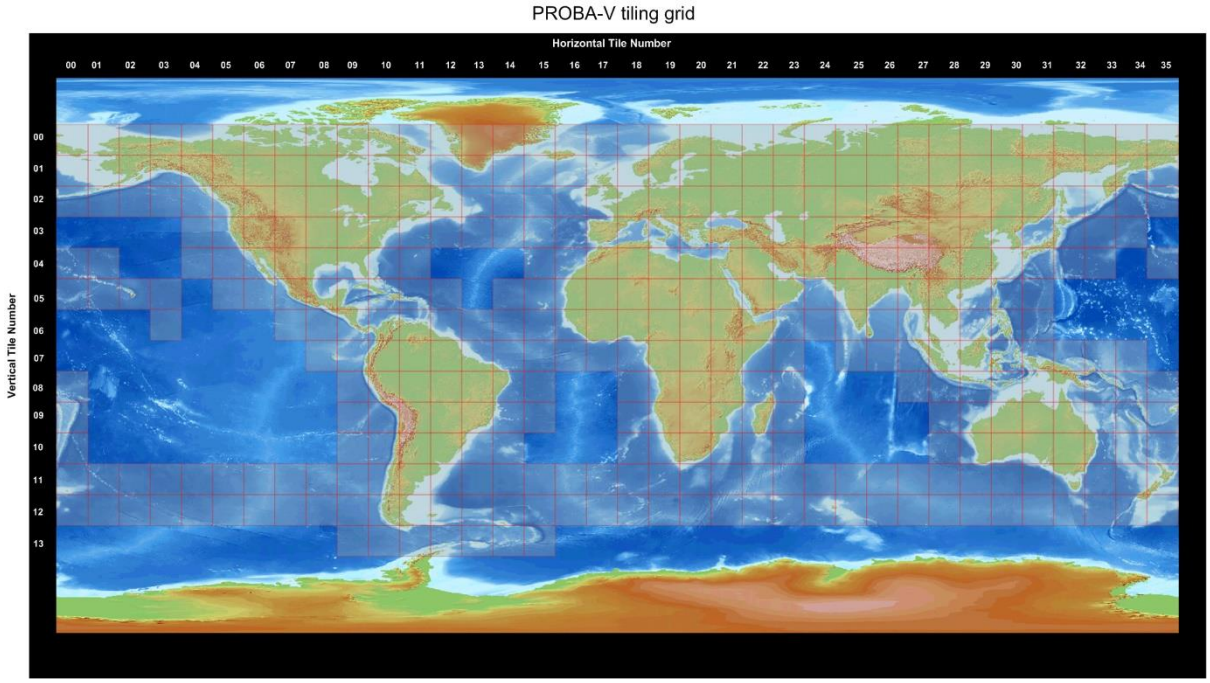

Figure 6: Grid tile definition for Proba-V and VGT.

5 Atmospheric correction

5.1 Computation of TOA reflectance and uncertainty

Note: In this section, among all the variables used in the AC algorithm, those that are constant (configuration parameters or inputs) are written in `Courier New` *font (*`input1`*), those that are computed (intermediate variable or outputs) are written in Math font (Δx).*

First, let's introduce the TOA reflectance (Eq. 3) and the associated uncertainty (Eq. 4) :

$$R_{toa}(\theta_s, \theta_v, \Delta\Phi) = \frac{\pi\, L(\theta_s, \theta_v, \Delta\Phi)}{\cos(\theta_s) E_s} \qquad Eq.\ 3$$

$$\Delta R_{toa}(\theta_s, \theta_v, \Delta\Phi) = \frac{\pi\, \Delta L(\theta_s, \theta_v, \Delta\Phi)}{\cos(\theta_s) E_s} \qquad Eq.\ 4$$

Where $L(\theta_s, \theta_v, \Delta\Phi)$ and $\Delta L(\theta_s, \theta_v, \Delta\Phi)$ `TOA-r` and `DTOA-r` respectively, `Es` are is the extra-terrestrial solar irradiance and θ_s is `SZA` .

5.2 Atmospheric correction method

5.2.1 Rationale

The Simplified Method for Atmospheric Corrections (Raman and Dedieu, 1994), SMAC, is based on the Second Simulation of the Satellite Signal in the Solar Spectrum, 6S, radiative transfer formulation of the satellite signal (Vermote et al., 1997) where all the pertinent radiative quantities are parameterized as a function of auxiliary data:

- Gas composition (mainly ozone and water vapour content for the spectral channels considered in this project)
- Aerosol content and aerosol type
- Molecular scattering mainly driven by the sea-level surface pressure and the surface elevation.

The choice of the SMAC algorithm is supported by the following arguments:

- It is operational and largely used in the land community and already implemented in the Copernicus Global Land Service (CGLOPS) and Copernicus Climate Change Service (C3S) processing lines.
- It is a robust and generic algorithm; thus it minimizes the dependence on the sensor which is beneficial for generating a multi sensor long time series with limited biases.
- The formulation of the algorithm is analytical and is adapted to an error propagation analysis.

5.2.2 Overview

The atmospheric correction is performed for <u>each band separately</u> and the basic equations of SMAC are:

$$T = T_g(\theta_s, \theta_v)\ T_{sca}(\theta_s)\, T_{sca}(\theta_v) = T_g(\theta_s, \theta_v)\, T_{atm} \qquad Eq.\ 5$$

$$R = R_{toa} - R_{atm}(\theta_s, \theta_v, \Delta\Phi)\ T_g(\theta_s, \theta_v) \qquad Eq.\ 6$$

$$R_{toc}(\theta_s, \theta_v, \Delta\Phi) = \frac{R}{T + s_{atm} R} \qquad Eq.\ 7$$

where the inputs are the geometry (`SZA` θ_s, `VZA` θ_v, and `RAA` $\Delta\Phi$), the pixel geolocation, the TOA reflectances R_{toa} and the outputs are the TOC directional reflectances R_{toc}. The auxiliary data are used to compute the total gaseous transmission T_g, the atmospheric path radiance R_{atm}, the atmospheric spherical albedo s_{atm} and the total downward and upward scattering transmissions T_{sca}. These computations are done for each sensor's channel according to the SMAC parameterization and for one aerosol model. From Eq. 5 to Eq. 7, it is possible to derive the atmospherically corrected TOC directional reflectance uncertainty ΔR_{toc} from the uncertainties on the TOA reflectance and on the atmospheric reflectances and transmittances, which are themselves derived from the uncertainties on the gaseous content, surface pressure aerosol load and type.

5.2.3 Error Propagation

5.2.3.1 TOA reflectance uncertainty

The radiometric uncertainty of the TOA reflectance is ΔR_{toa}. The resulting uncertainty on the TOC directional reflectance $\Delta_{R_{toc}}^{R_{toa}}$, is obtained by Eq. 8:

$$\Delta_{R_{toc}}^{R_{toa}} = \left|\frac{\partial R_{toc}}{\partial R_{toa}} . \Delta R_{toa}\right| \qquad Eq.\ 8$$

where $\frac{\partial R_{toc}}{\partial R_{toa}}$ is the sensitivity of TOC directional reflectance to any change in TOA reflectance. It is the Jacobian $J_{R_{toc}}^{R_{toa}}$.

Within SMAC, this Jacobian can be derived analytically (Eq. 9):

$$J_{R_{toc}}^{R_{toa}} = \frac{\partial R_{toc}}{\partial R} \frac{\partial R}{\partial R_{toa}} \qquad Eq.\ 9$$

after remarking that $\frac{\partial R}{\partial R_{toa}} = 1$, and using Eq. 10 in Eq 9:

$$\eta = 1/(T + s_{atm}\, R) \qquad Eq.\ 10$$

One obtains:

$$J_{R_{toc}}^{R_{toa}} = \eta^2 T \qquad Eq.\ 11$$

With atmospheric spherical albedo s_{atm}, *T* defined in Eq. 5 and R defined in Eq. 6.

5.2.3.2 Auxiliary data uncertainty

Ozone and Water Vapour

The sensitivity of TOC directional reflectance to the uncertainty in the total column of either ozone or water vapor is also treated analytically within SMAC. In the following, we apply the equation for the gas X, with X being O_3 or H_2O.

The uncertainty in the total column U_X is propagated through the Jacobian $J_{R_{toc}}^{U_X}$ (Eq. 12):

$$\Delta_{R_{toc}}^{U_X} = \left|J_{R_{toc}}^{U_X} . \Delta U_X\right| \qquad Eq.\ 12$$

Within SMAC, it can also be derived analytically.

Let us recall the formulation of the transmission of the gas T_X:

$$T_X = e^{\,a_X (U_X m)^{n_X}} \qquad Eq.\ 13$$

where a_X and n_X are SMAC coefficients and m is the air mass:

$$m = \frac{1}{\cos\theta_s} + \frac{1}{\cos\theta_v} \qquad Eq.\ 14$$

We decompose the Jacobian into two parts:

$$J_{R_{toc}}^{U_X} = \frac{\partial R_{toc}}{\partial T_X} . \frac{\partial T_X}{\partial U_X} \qquad Eq.\ 15$$

The sensitivity of T_X to U_X is:

$$\frac{\partial T_X}{\partial U_X} = \left(\frac{a_X . n_X}{U_X}\right) . (U_X m)^{n_X} . T_X \qquad \textit{Eq. 16}$$

The sensitivity of TOC directional reflectance to T_X is obtained through Eq. 13 to Eq. 16 and using the decomposition $T_g = T_X \, . T_{noX}$:

$$\frac{\partial R_{toc}}{\partial T_X} = -\eta^2 \frac{TR_{toa}}{T_X} \qquad \textit{Eq. 17}$$

Surface Pressure and AOT at 550 nm

For the surface pressure P_s and the AOT at 550 nm τ_a^{550}, we process the pixel two times more with small perturbations δP_s and $\delta\tau_a^{550}$, and then derive the Jacobians using the finite differences:

$$J_{R_{toc}}^{P_s} = \frac{\mathrm{R}_{toc}(P_s) - \rho_{toc}(P_s - \delta P_s)}{\delta P_s}, \qquad \textit{Eq. 18}$$

$$\Delta_{R_{toc}}^{P_s} = \left| J_{R_{toc}}^{P_s} . \Delta P_s \right|, \qquad \textit{Eq. 19}$$

$$J_{R_{toc}}^{\tau_a^{550}} = \frac{R_{toc}(\tau_a^{550}) - R_{toc}(\tau_a^{550} - \delta\tau_a^{550})}{\delta\tau_a^{550}} \qquad \textit{Eq. 20}$$

$$\Delta_{R_{toc}}^{\tau_a^{550}} = \left| J_{R_{toc}}^{\tau_a^{550}} . \Delta\tau_a^{550} \right| \qquad \textit{Eq. 21}$$

We use $\delta P_s = 10\ hPa, \delta\tau_a^{550} = 0.1\ \tau_a^{550}$. These values are chosen for provoking a sufficient response in the AC scheme and still staying in the linear regime.

Combination of all uncertainties

The individual errors are supposed independent and Gaussian. They are combined quadratically in Eq. 22:

$$\Delta R_{toc} = \sqrt{\left[\left(\Delta_{R_{toc}}^{R_{toa}}\right)^2 + \left(\Delta_{R_{toc}}^{U_{H_2O}}\right)^2 + \left(\Delta_{R_{toc}}^{U_{O_3}}\right)^2 + \left(\Delta_{R_{toc}}^{P_s}\right)^2 + \left(\Delta_{R_{toc}}^{\tau_a^{550}}\right)^2\right]} \qquad \textit{Eq. 22}$$

5.2.4 SMAC coefficients

These coefficients constitute the core of the SMAC algorithm. They have been calculated for all sensors and for (i) 148 aerosol models for VGT, PROBA-V, AVHRR and VIIRS and (ii) for only one aerosol model, `Continental`, for OLCI due to the near real time constraint of production within CGLOPS, of which the resulting TOC reflectances are used within this project to avoid processing too many datasets (budgetary constraint). The method of computation of the SMAC coefficients is described in Annex 1.

5.2.5 Errors characterization

5.2.5.1 TOA reflectance

The error on TOA radiance ΔL_{toa}comes from the inputs. If uncertainty on TOA reflectances is not furnished, it has been estimated, for example, for AVHRR sensors onboard Metop Platforms, VIIRS sensors and OLCI.
The TOA reflectance uncertainty is discussed in section 6, where the specific processing steps are described for each sensor separately.

5.2.5.2 Atmospheric and input parameters

Uncertainties on atmospheric parameters are described in detail (i) in [C3S_ATBD_CDR_SA_MULTI_SENSOR_v2.0] for VGT and PROBA-V sensors and (ii) in [CGLOPS_ATBD_S3-AC-V1.1] for OLCI sensors. Table 5 lists uncertainty on input parameters for AVHRR and VIIRS.

Table 5: Uncertainty of input parameters.

Uncertainty	Source	Value	Note
ΔR_{toa}	**From FCDR Level 1/pixel. *See section 6.1.3 for Metop-AVHRR and section 6.2.3 for VIIRS.***	**%**	
ΔU_{O_3}	Auxiliary data/statistical	6%	See Wargan et al., (2017), Davis et al., (2017)
ΔU_{H_2O}	Auxiliary data/statistical	20%	
ΔP_0, $\Delta\lambda$	Auxiliary data/statistical	<1 hPa , 0.002 K.m^{-1}	See van den Besselaar et al., (2011)
Δgeo	From Level 1/ pixel	Not known	
Δz	From FCDR Level 1 and auxiliary data/pixel	= f (DEM, Δgeo) OR Auxiliary `Delev` parameter	Complex on the fly computation OR If 2D geolocation error is considered random and its std. dev. is fixed, it can be stored in LUT with the DEM (convolution).
ΔP_S	From FCDR Level 1 and auxiliary data/pixel	From ΔP_0, $\Delta\lambda$, Δz: ΔP_S^{λ}~1hPa	depending on altitude and temperature vertical profile, See van den Besselaar et al., (2011)
$\Delta\tau_a^{550}$	Auxiliary data/statistical	$0.05 + 0.15\tau_a^{550}$ (2000-present) $0.07 + 0.20\,\tau_a^{550}$ (1980-1999)	See Randles et al., (2017).

5.2.6 Implementation

The SMAC algorithm has been implemented in C and Python under the GPL license (see for example http://www.cesbio.ups-tlse.fr/fr/modeles/modeles_list_8.html). We have decided to move to parallel based processing because it is a mature technology now, and we intend to run the SMAC algorithm a large number of times for each pixel in the forthcoming versions of the AC procedure. For this purpose, a new version called SMAC-GPU has been developed with the OPEN-CL graphics card programming language that includes the error propagation model with a Python interface.

5.3 Output files definition

The output files after atmospheric correction contain layers with TOC values and corresponding uncertainties for all relevant bands. Additionally, there are four layers related to the solar and viewing angles, cloud flag masks and a layer with the atmospheric correction flags (ac_flag). The definition of the bits in this layer is listed in Table 6. A more detailed overview of the different layers is given in Section 6.1.5 for Metop-AVHRR and in Section 6.2.6 for VIIRS.

Table 6: Characteristics of the atmospheric corrections flag (ac_flag) for AVHRR and VIIRS.

Bit	Name	Flag value
0	Reserved – not implemented	0 fixed
2,1	Bit 2: 0, Bit 1: 0 Nominal AOT value (AOT ≤ 0.5): High confidence level	0
	Bit 2: 0, Bit 1: 1 High AOT value (0.5 < AOT ≤1.0): Moderate confidence level	2
	Bit 2:1, Bit 1: 0 Very high AOT value (1.0 < AOT ≤1.5): Low confidence level	
	Bit 2:1, Bit 1: 1 Critical AOT value (1.5 < AOT): Very low confidence level	6
3	High solar zenith angle (SZA > 65°): Low confidence level	8
4	High viewing zenith angle (VZA > 65°): Low confidence level	16
5	Not used	32
6	Not used	64
7	Not used	128

6 Pre-processing description per sensor

In the following sections, the pre-processing is discussed per sensor to discuss sensor-specific approaches.

The SMAC algorithm is a robust and generic algorithm; thus it minimizes the dependence on the sensor which is an asset when the aim is a multi-sensor long time series with limited biases. Moreover, the SMAC algorithm is analytically sensor band independent.
Only for Sentinel-3 OLCI the processing of the atmospheric correction differs from that of the other sensors. For OLCI the aerosol model is assumed to be continental whereas for other sensors, the aerosol model is chosen among 148 models using MERRA-2 AOT fraction from 5 components (see section 3.2.3).

6.1 Metop-AVHRR

6.1.1 General approach

Figure 7 outlines the pre-processing approach for Metop-AVHRR.

Figure 7: Pre-processing outline for Metop-AVHRR.

6.1.2 Input data description

The Metop-AVHRR Level 1B products contain radiance and reflectance layers for all channels. Here we only use the TOA reflectance values for channels 1, 2, and 3a. The geolocation and angular information

are included as tie-point-grids with the same resolution as the channel layers. The input data also include a bit-wise cloud mask. The list of input layers is summarized in Table 7.

Table 7: Details of the relevant layers available in the Metop-AVHRR Level 1B products.

Layer	Size	Bit-type	Units	No data value
Radiances and reflectances				
radiance_N with N ∈ [1, 2, 3a, 3b]	Img	Int16	W / (m^2 sr)	0
reflec_N with N ∈ [1, 2, 3a, 3b]	Img	Int16	%	0
Geolocation information				
Latitude	Tie-point-grid	Float32	Degrees	0
Longitude	Tie-point-grid	Float32	Degrees	0
Angular information				
sun_zenith	Tie-point-grid	Float32	Degrees	0
sun_azimuth	Tie-point-grid	Float32	Degrees	0
view_zenith	Tie-point-grid	Float32	Degrees	0
view_azimuth	Tie-point-grid	Float32	Degrees	0
Quality flags				
Flags	Img	uInt8	-	0
cloud_flags	Img	uInt16	-	0

The cloud_flags layer is a bit-wise quality layer indicating the result of several test cases. The bit definition of this layer is given in Table 8. The test cases always have an ‘a’ and ‘b’ version, where the ‘a’ version is the test over land and the ‘b’ version is the test case over sea. They are used to detect different cloud types:

- The T4 test reveals low temperature corresponding to medium or high clouds.
- With the albedo tests, low clouds with a greater reflectivity than the underlying land (a) and the sea surface (b) can be detected.
- T4 – T5 detects cirrus clouds.
- T4 – T3 is applied to detect low-water clouds.
- T3 – T5 is used to detect semi-transparent ice clouds or sub-FOV cold clouds during night-time.
- The uniformity test is used to detect cloud edges, thin cirrus and small cumulus.

Table 8: Definition of the AVHRR cloud layer bits.

Bit	Value	Name	Description
0 – 3	1, 2, 4, 8	Number_of_the_test_situation	There are 11 different test situations
4	16	T4_test_b_clear_or_snow_ice	0 = test failed or cloudy 1 = clear or snow/ice covered
5	32	T4_test_a_cloudy_or_snow_ice	0 = test failed or clear 1 = cloudy or snow/ice covered
6	64	Albedo_test_b_clear_or_snow_ice	0 = test failed or cloudy 1 = clear or snow/ice covered
7	128	Albedo_test_a_cloudy_or_snow_ice	0 = test failed or clear 1 = cloudy or snow/ice covered
8	256	T4-T5_test_b_clear	0 = test failed or cloudy 1 = clear
9	512	T4-T5_test_a_cloudy	0 = test failed or clear 1 = cloudy

10	1024	T4-T3_test_b_clear	0 = test failed or cloudy 1 = clear
11	2048	T4-T3_test_a_cloudy	0 = test failed or clear 1 = cloudy
12	4096	T3-T5_test_b_clear	0 = test failed or cloudy 1 = clear
13	8192	T3-T5_test_a_cloudy	0 = test failed or clear 1 = cloudy
14	16384	Uniformity_test_b_clear	0 = test failed or cloudy 1 = clear
15	32768	Uniformity_test_a_cloudy	0 = test failed or clear 1 = cloudy

6.1.3 Assumed TOA uncertainties

The Metop-A, B, C L1B TOA reflectance products do not contain a specification of uncertainty. In order to get an estimate of this quantity, we borrow from the considerations made by the FIDUCEO team for the AVHRR/3 sensor (Mittaz et al. 2019), which are in turn based on data from the PATMOS-x project (https://cimss.ssec.wisc.edu/patmosx/calibration/) and literature review.

The TOA reflectance uncertainty is split into three independent contributions for each sensor: independent, structured, and common uncertainty. The independent uncertainty is assumed to be uncorrelated spatially and temporarily (e.g. caused by detector noise). For Metop-A, it is given in the FIDUCEO files as absolute value, which has some minor variation over the years. The structured uncertainty is correlated (e.g. between sensor autocalibrations). It is also given as absolute value with small variations. The common uncertainty is correlated over all measurements and is caused by imperfect knowledge of the spectral response functions and calibration coefficients. The latter has been estimated by the FIDUCEO team from literature review only and is given as a fraction of the actual reflectance.

The TOA reflectance uncertainty for channel $X \in \{1, 2, 3a\}$, δTOA_X, of a single pixel therefore is:

$$\delta TOA_X = \sqrt{(\delta independent_X)^2 + (\delta structured_X)^2 + (\delta common_X * TOA_X)^2}. \qquad \textit{Eq. 23}$$

TOA_X is the TOA reflectance value in channel X, while for the δ-terms the values from Table 9 are suggested. Note that the value for $\delta common_X$ in the files available through CEDA is given in fractions of 1, in spite of the specification of "%" in the metadata of the files (Mittaz, personal communication by email 2024-05-08).

Table 9: Uncertainty ranges from PUG of Mittaz et al. (2019).

Metop-A Channel	$\boldsymbol{\delta independent_X}$		$\boldsymbol{\delta structured_X}$ **(absolute)**		$\boldsymbol{\delta common_X}$ **(relative)**	
	min	max	min	max	min	max
1	0	2	0.001295	0.013819	0.03	0.03
2	0	2	0.001047	0.012615	0.05	0.05
3a	0	2	0.000737	0.024017	0.05	0.05

The correlations within a scanline (structured) or over the whole product (common) start to matter when multiple pixel uncertainties enter a computation, e.g. in an area average. Correlations of uncertainty across scanlines and channels are also reported by FIDUCEO. However, in the channel correlation matrices given in the dataset, significant correlations are only given for channels 5 and 6 (both thermal IR). See also Wooliams et al (2017) and https://research.reading.ac.uk/fiduceo/wp-

content/uploads/sites/129/2020/12/ErrorCovFormulaeEasyFCDR-v1.a.pdf for more details on the concepts of FIDUCEO uncertainty information and the importance of correlations of uncertainties.

Given that the FIDUCEO METOP-A dataset is not the immediate input to the atmospheric correction (and cannot be, because of an error in the channel 3a reflectance), the pragmatic approach to the uncertainty quantification adopted here is to use the values from the columns "mean" and "constant" of Table 10 in $\delta TOA_X = \sqrt{(\delta independent_X)^2 + (\delta structured_X)^2 + (\delta common_X * TOA_X)^2}$. Eq. 23. These values were obtained by inspecting a small sample of the FIDUCEO Metop-A files and filtering for pixels without quality issues and a solar zenith angle less than 60 degrees.

Table 10: Uncertainty ranges found in a sample of FIDUCEO Metop-A files filtered with SZA<60 and quality_pixel_bitmask==0x0).

Metop-A Channel	$\delta independent_X$ (absolute)			$\delta structured_X$ (absolute)			$\delta common_X$ (relative)
	min	**mean**	max	min	**mean**	max	**Constant**
1	0.0041	0.0041	0.02	0.00132	0.00135	0.00145	0.03
2	0.00409	0.0041	0.02	0.00109	0.0011	0.00125	0.05
3a	0.02	0.02	0.02	0.00076	0.0008	0.00107	0.05

6.1.4 Sensor-specific projection and tiling

Data with VZA larger than 63° or SZA larger than 65° are omitted. The VZA constraint removes data with large spatial deformations. The SZA constraint is related to the reliability of the atmospheric correction. The data is removed in the projection and tiling step.

6.1.5 Output file definition

The AVHRR output files contain:

- The TOC reflectances for the four solar channels (TOC_1, TOC_2, TOC_3a, and TOC_3b) and their associated uncertainties (TOC_1_error, TOC_2_error, TOC_3a_error, and TOC_3b error)
- The sun-view angles (SZA, VZA, SAA, and VAA)
- The distance to the nearest pixel in the source data, expressed as row offset, column offset, and in meters
- The cloud flag from the input file (see Table 8)
- The status map
- The confidence flag of the atmospheric corrections (see Table 6).

An overview of the different layers and their associated information is given in Table 11.

Table 11: Summary of AVHRR output layers.

Layer	**Bit-type**	**Units**	**Scale**	**Offset**	**Valid min**	**Valid max**	**Fill value**
TOC related layers							
TOC_N with N ∈ [1, 2, 3a, 3b]	Float32	-	5e-5	0	-31999	32767	-32000
TOC_N_error with N ∈ [1, 2, 3a, 3b]	Float32	-	5e-5	0	-31999	32767	-32000
Pixel distance related layers							
nncol	Int32	-	1	0	0	2047	-1

nnrow	Int32	-	1	0	0	1073	-1
nndist	Int32	meters	1	0	0	1556	-1
Angular information							
VAA	Float32	degrees	0.01	0	-31999	32767	-32000
VZA	Float32	degrees	0.01	0	-31999	32767	-32000
SAA	Float32	degrees	0.01	0	-31999	32767	-32000
SZA	Float32	degrees	0.01	0	-31999	32767	-32000
Cloud information							
cloud	Int32	-	1	0	0	65536	-1
status_map	uInt8	-	1	0	-	-	0
Atmospheric correction information							
ac_flag	Int32	-	1	0	0	-	-1

The meaning of the values in the status map are given in Table 12.

Table 12: Meaning of the AVHRR status map.

Value	Status map meaning
1	Clear
2	Ice_snow
4	Cloudy

6.2 VIIRS

6.2.1 General approach

Figure 8 outlines the pre-processing approach for VIIRS.

Figure 8: Pre-processing outline for VIIRS.

6.2.2 Input data description

There are various datasets available containing VIIRS data. In this project a combination of three datasets is used: VNP02MOD, VJ102MOD, and CLDMSK_L2_VIIRS_SNPP/NOAA.

6.2.2.1 VNP02MOD and VJ102MOD

VNP02MOD is the short-name for the SNPP-based NASA VIIRS L1B calibrated TOA radiances product derived from the sixteen M-bands. VJ102MOD is the analogue data for NOAA-20. These products are derived from the NASA VIIRS L1A raw radiances and include calibrated and geolocated radiance and reflectance data, quality flags, uncertainty indices, and an array of related information. The image dimensions of the 750 m swath product measure 3232 lines by 3200 pixels. Only the RSB bands M1 to M11 are used in CCI+ VP, the I-bands are not available in the MOD products.

The M-band data are stored as scaled integers. For the RSB bands, scale factor, offset, and unit attributes are provided for both reflectance and radiance values in Table 13. The stored reflectance is the product of the true reflectance and the cosine of the solar zenith angle at the pixel location. To obtain the reflectance, the reconstituted reflectance must be divided by the cosine of solar zenith angle. The radiance and reflectance offset values are both zero.

Information about the uncertainty indices and quality flags are given in Table 14 and Table 15. The meaning of the different quality flags is explained in Table 16.

More information about these products can be found here: https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/VNP02MOD .

Table 13: Details of L1B calibrated M-bands available in the VNP102MOD and VJP102MOD products (Table 9 in VIIRS L1B User Guide[2]).

M-bands 01 – 11 Earth-view reflectance variables

SDS layer	Bit-type	Radiance Units	Fill-value	Valid range	Radiance scale factor*	Scale factor	Flag values
M-band 01 EV reflectance	16-bit unsigned integer	watts/m²/steradian/µm	65535	Max = 65527 Min = 0	0.010990608	1.9991758E-5	65532, 65533, 65534**
M-band 02 EV reflectance	16-bit unsigned integer	watts/m²/steradian/µm	65535	Max = 65527 Min = 0	0.011957901	1.9991758E-5	65532, 65533, 65534**
M-band 03 EV reflectance	16-bit unsigned integer	watts/m²/steradian/µm	65535	Max = 65527 Min = 0	0.0128436815	1.9991758E-5	65532, 65533, 65534**
M-band 04 EV reflectance	16-bit unsigned integer	watts/m²/steradian/µm	65535	Max = 65527 Min = 0	0.012003864	1.9991758E-5	65532, 65533, 65534**
M-band 05 EV reflectance	16-bit unsigned integer	watts/m²/steradian/µm	65535	Max = 65527 Min = 0	0.0098244045	1.9991758E-5	65532, 65533, 65534**
M-band 06 EV reflectance	16-bit unsigned integer	watts/m²/steradian/µm	65535	Max = 65527 Min = 0	0.008220849	1.9991758E-5	65532, 65533, 65534**
M-band 07 EV reflectance	16-bit unsigned integer	watts/m²/steradian/µm	65535	Max = 65527 Min = 0	0.0061797476	1.9991758E-5	65532, 65533, 65534**
M-band 08 EV reflectance	16-bit unsigned integer	watts/m²/steradian/µm	65535	Max = 65527 Min = 0	0.002928277	1.9991758E-5	65532, 65533, 65534**
M-band 09 EV reflectance	16-bit unsigned integer	watts/m²/steradian/µm	65535	Max = 65527 Min = 0	0.0023088641	1.9991758E-5	65532, 65533, 65534**
M-band 10 EV reflectance	16-bit unsigned integer	watts/m²/steradian/µm	65535	Max = 65527 Min = 0	0.0015735304	1.9991758E-5	65532, 65533, 65534**
M-band 11 EV reflectance	16-bit unsigned integer	watts/m²/steradian/µm	65535	Max = 65527 Min = 0	4.8158638E-4	1.9991758E-5	65532, 65533, 65534**

* Radiance scale factor values are granule-dependent

** Flag meanings: 65532 = Missing_EV, 65533 = Bowtie_Deleted, 65534 = Cal_Fail

2 https://ladsweb.modaps.eosdis.nasa.gov/api/v2/content/archives/Document%20Archive/Science%20Data%20Product%20Documentation/NASA_VIIRS_L1B_UG_August_2021.pdf

Table 14: Information about the uncertainty indices available in the VNP102MOD and VJP102MOD products (from Table 9 in VIIRS L1B User Guide).

M-bands 01 – 16 Uncertainty indices (for 16 separate bands)

SDS layer	Bit-type	Units	Fill-value	Valid range	Conversion	Scale factor
M-band 01 – 16 uncertainty indices	16-bit integer	percent	-1	Max = 127 Min = 0	1.0 + scale * index^2	0.006138

Table 15: Information about the quality flag layer in the VNP102MOD and VJP102MOD products (from Table 9 in VIIRS L1B User Guide).

M-bands Quality flags

SDS layer	Bit-type	Flag masks
M-band 01 – 05, 07, 13 quality flags	16-bit unsigned integer	1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096*
M-band 06, 08 – 12, 14 – 16 quality flags	16-bit unsigned integer	1, 2, 4, 8, 256, 512, 1024, 2048, 4096**

* Flag meanings: 1 = Substitute_Cal, 2 = Out_of_Range, 4 = Saturation, 8 = Temp_not_Nominal, 16 = Low_Gain, 32 = Mixed_Gain, 64 = DG_Anomaly, 128 = Some_Saturation, 256 = Bowtie_Deleted, 512 = Missing_EV, 1024 = Cal_Fail, 2048 = Dead_Detector, 4096 = Noisy_Detector

** Flag meanings: 1 = Substitute_Cal, 2 = Out_of_Range, 4 = Saturation, 8 = Temp_not_Nominal, 256 = Bowtie_Deleted, 512 = Missing_EV, 1024 = Cal_Fail, 2048 = Dead_Detector, 4096 = Noisy_Detector

Table 16: Meaning of the M-band quality flags in the VNP102MOD and VJP102MOD products (from Table-C.1 in VIIRS L1B UG).

Value	Flag Meaning	Description
1 (LSB)	Substitute_Cal	Granule-average has been substituted for SV and/or BB
2[1]	Out_of_Range	Earth view counts &/or cal. radiance out of range (see Table A.3)
4[1]	Saturation	L1A Earth view counts ≥ 4095
8	Temp_not_Nominal	Measured temperatures outside nominal range
16[2]	Low_Gain	One or more low-gain samples in aggregation
32[2]	Mixed_Gain	Mix of high- and low-gain samples in aggregation
64[2]	DG_Anomaly	Dual-gain anomaly
128[2]	Some_Saturation	One or more samples in aggregation have L1A Earth view counts that are either saturated or out of range
256	Bowtie_Deleted	Data excluded by VIIRS for bowtie deletion
512	Missing_EV	Packet missing or corrupted in transmission
1024	Cal_Fail	Calibration failure
2048	Dead_Detector	Detector is not producing valid data
4096	Noisy_Detector	Noisy Detector

[1] The criteria for assigning Saturation and Out_of_Range pixel quality flags are:

- If a pixel raw DN value equals or exceeds the DN limit of 4095, the pixel "Saturation" quality flag is set, and the DN value is set to 4095.
- If a pixel raw DN value exceeds the "Out-of-Range" value, the pixel "Out of Range" quality flag is set.

[2] Dual-gain M-bands only. Application of Low-Gain and Mixed-Gain pixel flags is as follows:

- One sample aggregation: high-gain 00; low-gain 01
- Two sample aggregation: all high-gain 00; all low-gain 01; mixed 11
- Three sample aggregation: all high-gain 00; all low-gain 01; two low-gain 11; two high-gain 10

6.2.2.2 CLDMSK_L2_VIIRS_SNPP / NOAA20

CLDMSK_L2_VIIRS_SNPP / NOAA20 is a cloud mask dataset that can be used as an alternative source for the status map. It is derived from the MODIS-VIIRS Cloud Mask (MVCM) algorithm which is one of the continuity algorithms to facilitate continuity between Moderate Resolution Imaging Spectroradiometer (MODIS) and VIIRS data. In addition to the angular information (SAA, SZA, VAA, and VZA) and the geolocation information (longitude and latitude), CLDMSK_L2_VIIRS_SNPP / NOAA20 also contains an extensive cloud mask ('Cloud_mask') and quality assurance layer, and two other cloud mask output arrays. These last two arrays are the 'Clear_Sky_Confidence' and 'Integer_Cloud_Mask' data sets.

The 'Clear_Sky_Confidence' is the final numeric value of the confidence of clear sky, or Q value (Ackerman et al., 2010), computed by the MVCM algorithm. It is this value that is converted into one of four cloud mask categories (confident clear, probably clear, probably cloudy, confident cloudy) reported in bits 1 and 2 (0-based) of the 'Cloud_mask', but this is not used in CCI+ VP.

The other array, 'Integer_Cloud_Mask', is the value of bits 1 and 2 of the cloud mask converted to an integer value. This is for users who would like to use the cloud mask categories but with no need to unpack the remaining bits in the cloud mask array. The integer values are 0-3, corresponding to confident cloudy, probably cloudy, probably clear, and confident clear, respectively. A value of -1

indicates no result (fill value). This data set does not include information on snow/ice, as the snow/ice background flag refers to if an ancillary dataset on snow/ice presence was used in the cloud test. It does not give information on the presence of snow/ice.

6.2.3 Assumed TOA uncertainties

The products VNP02MOD and VJ102MOD contain a layer for TOA reflectance uncertainty, which is filled with a constant value differing from the netCDF "_FillValue". Unfortunately, this layer is only a placeholder (modapsuso@lists.nasa.gov, ticket #138799, personal communication, 2024-05-16). Two publications on the subject were pointed out in this communication. Sun et al. (2021) shows in their Figure 19 calibration uncertainties of SNPP VIIRS reflective solar bands and quantify the in-orbit detector signal-to-noise ratio. However, there seem to be other sources of uncertainty, which are dominant, at least in the VIS/NIR bands. Twedt et al. (2022) found differences between the NOAA-20 VIIRS and SNPP L1B reflectance bands of 3–6% for the VIS/NIR bands and 2–3% for the SWIR. These differences are stable in time through the first three years of the NOAA-20 mission, indicating a static calibration error.

Comparing NOAA-20 VIIRS and SNPP directly, Moyer et al. (2021) quantified the systematic differences between the sensors and identified a persistent calibration issue of SNPP VIIRS as the most probable cause. The cross-calibration factors are used to quantify an additional source of calibration uncertainty:

$$\delta TOA = \sqrt{(bias + trend * \Delta t)^2 + (\max(\delta\mathrm{SNPP}, \delta\mathrm{NOAA\text{-}20}))^2}, \qquad \text{(Eq. 1)}$$

with δTOA the cross-calibrated TOA uncertainty, *bias* and *trend* are the mean bias and trend of the mean bias between SNPP and NOAA-20 VIIRS reflectances from Figure 3 in Moyer et al. (2021). Δt is the period in years over which the bias and trend is taken. $\delta\mathrm{SNPP}$ is the SNPP calibration uncertainty from Figure 19 in Sun et al. (2021) and $\delta\mathrm{NOAA\text{-}20}$ is the NOAA-20 TOA reflectance accuracy from Table 5 in Moyer et al. (2021).

All these values are summarized for the RSB M-bands in Table 17. The second and third column are the mean bias and mean bias trend from Moyer et al. (2021). The calibrated SNPP TOA uncertainties from Sun et al. (2021) are given in the fourth column, while the NOAA-20 reflectance accuracies for the low gain bands from Moyer et al. (2021) are summarized in the fifth column. The sixth and seventh column contain the cross-calibrated TOA uncertainties for the beginning and end of the analyzed NOAA-20 period (Δt = 0 and 3 respectively), computed with $\delta TOA = \sqrt{(bias + trend * \Delta t)^2 + (\max(\delta\mathrm{SNPP}, \delta\mathrm{NOAA\text{-}20}))^2}$, (Eq. 1). As a further simplification, the maximum of the relative bias over the analysed period from Moyer et al. (2021) is used (last column of Table 17).

This formulation is merely a gross attempt at quantifying the uncertainty but may be sufficient for our purposes. It assumes that the measurements of the two instruments carry two independent realisations of the bias. A more rigorous mathematical attempt is complicated by the choice of a relative error model and would also require further idealising assumptions. The presented approach neglects among other things detector and quantisation noise, assuming that they are independent of and smaller than the calibration uncertainty. The choice of a purely relative uncertainty model without absolute components (cf. AVHRR), may be inappropriate for very low reflectances. Their quantification, however, would require a more detailed analysis than what can be performed within the current project.

VIIRS band M6 is excluded from most analyses, because of missing striping correction. Quoting from Sun et al. (2021): *"Band M6 is saturated at a lot of pixels when it views the Libya 4 site and the deep*

convective clouds (DCC). Thus, the striping correction coefficients have not been derived with the current methodology for band M6." Investigation of M6 TOA retrievals indicates that the striping introduces structured artifacts which have a magnitude of about 10% of the smoothed reflectance value. Together with an assumed calibration uncertainty of 6% (rounded maximum of the suggested values for other bands from Table 17), the uncertainty would be 12% (= $\sqrt{0.1^2 + 0.06^2}$). Band M6 may turn out not to be useful for retrievals and may be dropped at a later stage, also because of the low spatial coverage caused by the very frequent saturation of the band.

Table 17: Input data from Moyer et al. (2021) and Sun et al. (2021) for the computation of the time constant relative uncertainty of SNPP TOA uncertainties. See text for more explanation.

Band	Mean bias (%)	Mean bias trend (%/yr)	SNPP calibrated δTOA (%)	NOAA-20 δTOA low gain (%)	SNPP δTOA for Δyr = 0 (%)	SNPP δTOA for Δyr = 3 (%)	Final SNPP δTOA (%)
M1	-3.5	-0.03±0.10	1.2	1.9	3.98	4.06	**4.06**
M2	-1.9	0.06±0.10	1.2	1.78	2.60	2.48	**2.60**
M3	-2.6	0.02±0.11	1.2	1.88	3.21	3.16	**3.21**
M4	-2.8	0.01±0.11	1.2	1.41	3.13	3.11	**3.13**
M5	-4.4	-0.05±0.11	1.2	1.38	4.61	4.75	**4.75**
M6			1.2	1.5			**12**
M7	-3.5	0.11±0.09	1.2	1.43	3.78	3.48	**3.78**
M8	-2.8	-0.2±0.08	1.2	1.76	3.31	3.83	**3.83**
M9	-1.6	-0.15±0.18	1.2	1.48	2.18	2.53	**2.53**
M10	-3.4	-0.21±0.21	1.1	1.37	3.67	4.26	**4.26**
M11	-2.5	-0.17±0.16	5	2.25	5.59	5.84	**5.84**

6.2.4 Sensor-specific projection and tiling

Pixels marked with the 'Bowtie_Deleted' quality flag in the observation data are discarded for projection and tiling.

6.2.5 Sensor-specific atmospheric correction description

The atmospheric correction is performed as described in section 6.

6.2.6 Output file definition

THE VIIRS output files contain:

- The TOC reflectances for the 11 solar bands (TOC_N, with N from M1 to M11) and their associated uncertainties (TOC_N_error)
- The input quality flag associated to each band (TOC_N_quality_flags)
- The sun-view angles (solar azimuth, solar zenith, sensor azimuth, and sensor zenith)
- The cloud mask (Inter_Cloud_Mask) from the input file
- The clear-sky confidence flag (Clear_Sky_Confidence) from the input file
- The confidence flag of the atmospheric corrections (see Table 5).

The geolocation information from CLDMSK_L2_VIIRS_SNPP / NOAA20 is included implicitly in each layer. An overview of the different layers and their associated information is given in Table 18.

Table 18: Summary of VIIRS output layers. The origin file numbers correspond to 1 = VNP02MOD/VJ102MOD and 2 = CLDMSK_L2_VIIRS_SNPP / NOAA20.

Layer	Origin file	Bit-type	Units	Scale	Offset	Valid min	Valid max	Fill value
TOC related layers								
TOC_N with N from M1–M11	1	Float32	-	5e-5	0	-	-	-32000
TOC_N_error with N from M1–M11	1	Float32	-	5e-5	0	-	-	-32000
TOC_N_quality_flags with N from M1–M11	1	uInt16	-	1	0	-	-	65535
Angular information								
sensor_azimuth	2	Float32	degrees	0.01	0	-18000	18000	-32768
sensor_zenith	2	Float32	degrees	0.01	0	0	18000	-32768
solar_azimuth	2	Float32	degrees	0.01	0	-18000	18000	-32768
solar_zenith	2	Float32	degrees	0.01	0	0	18000	-32768
Cloud information								
Integer_Cloud_Mask	2	Int8	-	1	0	0	3	-1
Clear_Sky_Confidence	2	Float32	-	1	0	0	1	-999.9
Atmospheric correction information								
ac_flag	-	Int32	-	1	0	0	-	-1

In Table 19, the definition of the quality flags associated with the TOC data are given.

Table 19: Description of the TOC_N_quality_flags layer.

Value	Flag Meaning	Description
1	Substitute_Cal	Granule-average has been substituted for SV and/or BB
2	Out_of_Range	Earth view counts &/or cal. radiance out of range (see Table A.3)
4	Saturation	L1A Earth view counts ≥ 4095
8	Temp_not_Nominal	Measured temperatures outside nominal range
16	Low_Gain	One or more low-gain samples in aggregation
32	Mixed_Gain	Mix of high- and low-gain samples in aggregation
64	DG_Anomaly	Dual-gain anomaly
128	Some_Saturation	One or more samples in aggregation have L1A Earth view counts that are either saturated or out of range
256	Bowtie_Deleted	Data excluded by VIIRS for bowtie deletion
512	Missing_EV	Packet missing or corrupted in transmission
1024	Cal_Fail	Calibration failure
2048	Dead_Detector	Detector is not producing valid data
4096	Noisy_Detector	Noisy detector

The Integer_Cloud_Mask can take four integer values as summarized in Table 20.

Table 20: Description of the Integer_Cloud_Mask values of VIIRS.

Value	Cloud mask meaning
0	Cloudy
1	Probably cloudy
2	Probably clear
3	Confident clear

6.3 Sentinel-3 OLCI

The TOC reflectance of Sentinel-3 generated in the Copernicus Land Monitoring Service are used as input for CCI+ VP. These data are gridded at 333 m. To be able to use this data set for the retrieval of LAI and fAPAR at 1 km, these data are first re-gridded to 1 km, considering the associated per-pixel quality information. The procedure of the re-gridding is described in this section.

6.3.1 Input data description

The input TOC reflectances used are those for the OLCI sensor coming from the Copernicus Global Land Service – Lot1 (CGLOPS1). The data are NetCDF files, and the layers are summarized in Table 21. A more comprehensive explanation of the input data is given in the PUM [ED4], and the relevant information is copied below.

Table 21: Sentinel-3 input data

Parameter	Description
Oa*xx*_toc (for *xx* in 2-12, 16-18, 21)	TOC reflectances for OLCI bands
Oa*xx*_toc_error (for *xx* in 2-12, 16-18, 21)	Error on TOC reflectances for OLCI bands
SAA_OLCI	Solar azimuth angle for OLCI
SZA_OLCI	Solar zenith angle for OLCI
VAA_OLCI	Viewing azimuth angle for OLCI
VZA_OLCI	Viewing zenith angle for OLCI
Quality_flags	Quality flags from OLCI level 1b input data (see Table 22)
Pixel_classif_flags	Output from OLCI-IdePix input data (see Table 23)
AC_process_flag	Quality flag of the atmospheric correction processing (Table 6)
Latitude	Latitude
Longitude	Longitude

Table 22: Quality_flags encoding

Bit	Name	Value
0	saturated_Oa21	1
1	saturated_Oa20	2
2	saturated_Oa19	4
3	saturated_Oa18	8
4	saturated_Oa17	16
5	saturated_Oa16	32
6	saturated_Oa15	64
7	saturated_Oa14	128
8	saturated_Oa13	256
9	saturated_Oa12	512
10	saturated_Oa11	1024
11	saturated_Oa10	2048
12	saturated_Oa09	4096
13	saturated_Oa08	8192
14	saturated_Oa07	16384
15	saturated_Oa06	32768
16	saturated_Oa05	65536
17	saturated_Oa04	131072
18	saturated_Oa03	262144
19	saturated_Oa02	524288

20	saturated_Oa01	1048576
21	dubious	2097152
22	sun_glint_risk	4194304
23	duplicated	8388608
24	cosmetic	16777216
25	invalid	33554432
26	straylight_risk	67108864
27	bright	134217728
28	tidal_region	268435456
29	fresh_inland_water	536870912
30	coastline	1073741824
31	land	2147483648

Table 23: pixel_classif_flags encoding

Bit	Name	Value
0	IDEPIX_INVALID	1
1	IDEPIX_CLOUD	2
2	IDEPIX_CLOUD_AMBIGUOUS	4
3	IDEPIX_CLOUD_SURE	8
4	IDEPIX_CLOUD_BUFFER	16
5	IDEPIX_CLOUD_SHADOW	32
6	IDEPIX_SNOW_ICE	64
7	IDEPIX_BRIGHT	128
8	IDEPIX_WHITE	256
9	IDEPIX_COASTLINE	512
10	IDEPIX_LAND	1024
11	IDEPIX_MOUNTAIN_SHADOW	2048
DN no data value		**-1**

The A/C processing flags are defined as in Table 6.

6.3.2 Re-gridding to 1 km

6.3.2.1 Grid

The S3 333 m data should be re-gridded to exactly the same grid as PROBA-V 1 km data (used as reference):

- The grid uses the World Geodetic System 1984 (WSG 84) projection as defined by the following Well-Known Text (WKT):

```
GEOGCS["WGS                                                    84",
        DATUM["WGS_1984",
              PHEROID["WGS                                     84",
                      6378137,
                      298.257223563,
                      'AUTHORITY["EPSG","7030"]],
              TOWGS84[0,0,0,0,0,0,0],
              AUTHORITY["EPSG","6326"]],
        PRIMEM["Greenwich",
               0,
```

```
        AUTHORITY["EPSG","8901"]],
    UNIT["degree",
         0.0174532925199433,
         AUTHORITY["EPSG","9108"]],
    AUTHORITY["EPSG","4326"]]
```

- All coordinates are centre pixel coordinates. The bounding box coordinates (e.g. to be used by GDAL) can be calculated as:
  - $Y_{upper} = Y_{centerpixel} + \frac{1}{112}$ and $Y_{lower} = Y_{centerpixel} - \frac{1}{112}$
  - $X_{X_{centerpixel} - \frac{1}{112}}$ and $X = X_{centerpixel} + \frac{1}{112}$
- The upper left tie point of the grid is 75°N,180°W and covers the area up to 65°S,180°E.
- Individual tiles in the Grid have a tile index in their names (e.g., X10Y04), to represent the tiles relative horizontal (X) and vertical (Y) position on the globe. Note that the original PROBA-V grid also uses a XY index in their tile names. However, the S3 is extended 10° North and as such, the Y indexes have an offset of 1, e.g., PROBA-V tile X10Y04 corresponds with a Sentinel-3 tile X10Y05.

This means that tiles with the same upper left coordinate, but different resolutions, don't have the same bounding box coordinates.
Illustrated in Figure 9 are the coordinates for Sentinel3-grid 10°x10° tile X18Y03, both in 333 m and 1 km pixel resolutions.

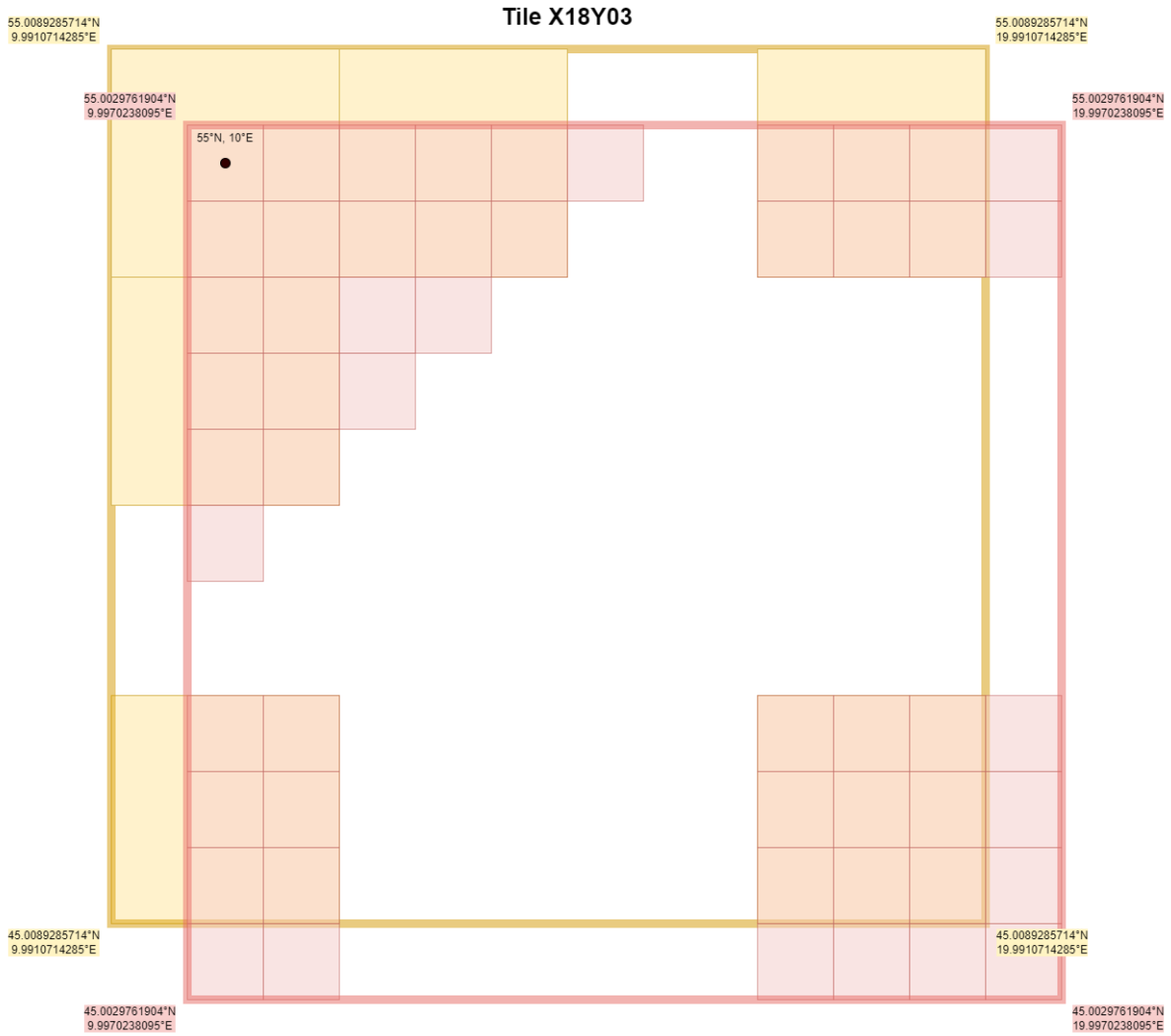

Figure 9: Illustration of the coordinates for Sentinel3-grid 10°x10° tile X18Y03, both in 333 m and 1 km pixel resolutions.

6.3.2.2 Selection of pixels

The aggregation of 3x3 pixels of the different layers will be performed only on pixels that meet certain conditions. We follow the recommendations on using the annotation flags as defined in CGLOPS1 (see [ED4]).

For OLCI reflectance bands, pixels with the following flags should be omitted:

- quality_flags layer (See Table 9 in [ED4]):
  - 'saturated_Oa*' depending on which band is used, exclude when raised.
  - 'sea
- pixel_classif_flags (See Table 10 in [ED4]):
  - IDEPIX_LAND: include when raised
  - IDEPIX_MOUNTAIN_SHADOW: include when raised
  - IDEPIX_INVALID: exclude when raised
  - IDEPIX_CLOUD: exclude when raised
  - IDEPIX_CLOUD_AMBIGUOUS: exclude when raised
  - IDEPIX_CLOUD_BUFFER: exclude when raised
  - IDEPIX_CLOUD_SHADOW: exclude when raised

- AC_process_flag (See Table 11 in [ED4]):
    - Exclude when AOT > 1 (Bit 2 set)
    - Exclude when SZA > 65° (Bit 3 set)

For all layers, only if 5 or more pixels are not flagged, then the values are resampled to 1 km. If only 4 or less pixels are not flagged, then the output value will be flagged (see section 6.3.2.3.4).

6.3.2.3 Resampling method per layer

6.3.2.3.1 OLCI Reflectances

Reflectances are averaged per band. Only pixels that are not flagged (see 6.3.2.2) are averaged. Further distinction should be made between clear land pixels and snow/ice land pixels according to the pixel_classif_flag:

- IDEPIX_SNOW_ICE

These two classes should never be mixed in the average. Only the pixels where the majority is (not snow/ice) or (snow/ice) should be averaged. The following rule should be applied:
After elimination of other flagged values, take the average of IDEPIX_LAND pixels if

- At least 5 pixels are retained after elimination of other flags, AND
- The majority of these pixels is NOT labelled as IDEPIX_SNOW_ICE, AND
- At least 4 pixels with IDEPIX_LAND are retained.
- The output flag is then set to LAND.

A similar rule should be applied for snow/ice, take the average of IDEPIX_SNOW_ICE pixels if:

- At least 5 pixels are retained after elimination of other flags, AND
- The majority of these pixels is labelled as IDEPIX_SNOW_ICE, AND
- At least 4 pixels are IDEPIX_SNOW_ICE.
- The output flag is then set to LAND and SNOW_ICE.

If there are less than 4 pixels either IDEPIX_SNOW_ICE, then

- the average of all these non-flagged pixels will be taken
- the output flag is set to MIXED SNOW/ICE/LAND

$$Oaxx_{toc\,g,h,1000} = \frac{1}{N}\sum_{n=1}^{N} Oaxx_{toc\,i,j,333}$$

with

N: ranging between 4 and 9.
xx: band
1000, 333: spatial resolution
i,j: pixel position in 333 m
g,h: pixel position in 1000 m

6.3.2.3.2 OLCI angles

The value of the middle pixel will be retained.

6.3.2.3.3 Uncertainty of TOC reflectance

Uncertainties are propagated only for the pixels selected for the TOC reflectance averaging (see sections 6.3.2.2 and 6.3.2.3.1).
The method to calculate the output uncertainty is, assuming that the uncertainties are uncorrelated:

$$Oaxx_{g,h,1000} = \frac{1}{N}\sqrt{\sum_{n=1}^{N} Oaxx_{i,j,333}^{2}}$$

with

N: ranging between 4 and 9.
xx: band
1000, 333: spatial resolution
i,j: pixel position in 333 m
g,h: pixel position in 1000 m

6.3.2.3.4 Output Flags

Output flags will be summarized in one quality layer as defined in Table 24.
Multiple alternative reasons exist why pixels within the 3x3 window are flagged and reflectance not calculated. Although it would be possible to propagate this information e.g. in the middle of large clouds because the reason for omission of the reflectance is the same for all pixels averaged, mostly this will inevitably result in multiple reasons why TOC reflectance was not calculated at 1000 m and these are not easy to label. In any case, it would not improve the retrievals.
For cases where the 1km OLCI TOC is calculated, some additional quality flags are available to identify possible quality issues in the 1km data. See table below.

Table 24: Output flag definitions.

Bit	Name	Description	Value
0	LAND	from IDEPIX_LAND	1
1	SNOW/ICE (1)	(Majority of good pixels are IDEPIX_SNOW) AND (at least 4 non-flagged pixels with IDEPIX_SNOW)	2
2	MIXED CLEAR/SNOW/ICE (1)	(Majority of good pixels are IDEPIX_SNOW_ICE or not IDEPIX_SNOW_ICE) AND (less than 4 non-flagged pixels with IDEPIX_SNOW_ICE or not IDEPIX_SNOW_ICE)	4
3	BRIGHT	(TOC reflectance is calculated) AND (at least 1 non-flagged pixel with IDEPIX_BRIGHT is used)	8
4	WHITE	(TOC reflectance is calculated) AND (at least 1 non-flagged pixel with IDEPIX_WHITE is used)	16
5	highAOT	A**t least 1** of the input pixels has flag 0.5<AOT<1.0 (see table 5)	32
6	highAOTall	A**ll** of the input pixels has flag 0.5<AOT<1.0 (see table 5)	64
7	MISSING (1)	TOC reflectance is not calculated	126

6.3.2.3.5 Latitude and longitude

Lat, lon will be determined by the grid specification, the 1km pixel will have the same lat/lon as the 3x3 center pixel.

6.3.3 Output layers

The output layers are summarized in Table 25. Data type and scaling are identical to the input data.

Table 25: Output layers definition

Parameter	Description
Oa*xx*_toc (for *xx* in 2-12, 16-18, 21)	TOC reflectances for OLCI bands
Oa*xx*_toc_error (for *xx* in 2-12, 16-18, 21)	Error on TOC reflectances for OLCI bands
SAA_OLCI	Solar azimuth angle for OLCI
SZA_OLCI	Solar zenith angle for OLCI
VAA_OLCI	Viewing azimuth angle for OLCI
VZA_OLCI	Viewing zenith angle for OLCI
Quality flag	Defined in Table 24
Latitude	Latitude
Longitude	Longitude

7 References

Ackerman, S. A., R. A. Frey, K. I. Strabala, Y. Liu, L. E. Gumley, B. A. Baum, and W. P. Menzel, Discriminating clear-sky from clouds with MODIS algorithm theoretical basis document (MOD35), Ver. 6.1, October 2010, MODIS Atmosphere Web Site, 117 pp., 2010

Gordon, I. E., L. S. Rothman, C. Hill, R. V. Kochanov, Y. Tan, *et al.*, (2017) The HITRAN2016 molecular spectroscopic database, *J. Quant. Spectrosc. Radiat. Transfer* **203**, 3-69 (2017)

Hess, M.; Koepke, P. & Schult, I, (1998). Optical Properties of Aerosols and Clouds: The Software Package OPAC. *Bulletin of the American Meteorological Society*, 1998, **79**, 831-844

Mittaz, J.; Taylor, M.; Desmons, M.; Mollard, J.; Merchant, C.J. (2019): FIDUCEO: Fundamental Climate Data Record of recalibrated brightness temperatures for the Advanced Very-High-Resolution Radiometer (AVHRR) with metrologically-traceable uncertainty estimates, 1998 - 2016, v1.0. Centre for Environmental Data Analysis, https://catalogue.ceda.ac.uk/uuid/67b206b17365444b8243182f04c8ba44

Moyer, D., S. Uprety, W. Wang, C. Cao, and I. Guch "S-NPP/NOAA-20 VIIRS reflective solar bands on-orbit calibration bias investigation", Proc. SPIE 11829, Earth Observing Systems XXVI, 1182912 (3 August 2021); https://doi.org/10.1117/12.2595175

R.V. Kochanov, I.E. Gordon, L.S. Rothman, P. Wcislo, C. Hill, J.S. Wilzewski (2016). HITRAN Application Programming Interface (HAPI): A comprehensive approach to working with spectroscopic data, *J. Quant. Spectrosc. Radiat. Transfer,* **177**, 15-30 (2016). https://hitran.org/hapi

Rahman H. and G. Dedieu, SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum. Intern. J. Remote Sens., 15, 123-143, 1994.

Schreier, F. and S. Gimeno Garcia (2013). Py4CAtS – Python Tools for Line-by-Line Modelling of Atmospheric Radiative Transfer. *In Robert F. Cahalan and Jürgen Fischer (editors), Radiation Processes in the Atmosphere and Ocean (IRS 2012): Proceedings of the International Radiation Symposium (IRC/IAMAS) Volume 1531 of AIP Conference Proceedings*, pages **123 - 126**. American Institute of Physics, 2013. doi: 10.1063/1.4804723 https://atmos.eoc.dlr.de/tools/Py4CAtS

Sun, Junqiang; Xiaoxiong Xiong, Ning Lei, Sherry Li, Kevin Twedt, and Amit Angal. 2021. "Ten Years of SNPP VIIRS Reflective Solar Bands On-Orbit Calibration and Performance" Remote Sensing 13, no. 15: 2944. https://doi.org/10.3390/rs13152944

Toté, C., E. Swinnen, E. Wolters and B. Smets (2024) Validation report: Normalized Difference Vegetation Index (ENDVI10), LSA-420, issue 2, LSA-SAF, https://nextcloud.lsasvcs.ipma.pt/s/zXJoeTf6HByE6RP?dir=undefined&path=%2FENDVI10_CDR_iCDR_V2%2FVR-Validation_Report&openfile=180234

Twedt, Kevin; Ning Lei, Xiaoxiong Xiong, Amit Angal, Sherry Li, Tiejun Chang, and Junqiang Sun. 2022, "On-Orbit Calibration and Performance of NOAA-20 VIIRS Reflective Solar Bands" IEEE Transactions on Geoscience and Remote Sensing 60, 1-13, https://doi.org/10.1109/TGRS.2021.3108970

Vermote, E.F., D. Tanré, J.L. Deuzé, M. Herman and J.-J. Morcrette, Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: An Overview, IEEE Transactions on Geoscience and Remote Sensing, Vol. 35, No. 3, p. 675-686, 1997.

Wooliams, E., Mittaz, J., Merchant, C., & Harris, P. (2017). FIDUCEO project: Principles behind the FCDR effects table. FIDUCEO project. https://doi.org/10.5281/zenodo.7357165

8 Annexes:

8.1 Annex 1: Computation of the SMAC coefficients

8.1.1 Processing chain

In Figure 10 is shown how SMAC coefficients are derived from basic data.

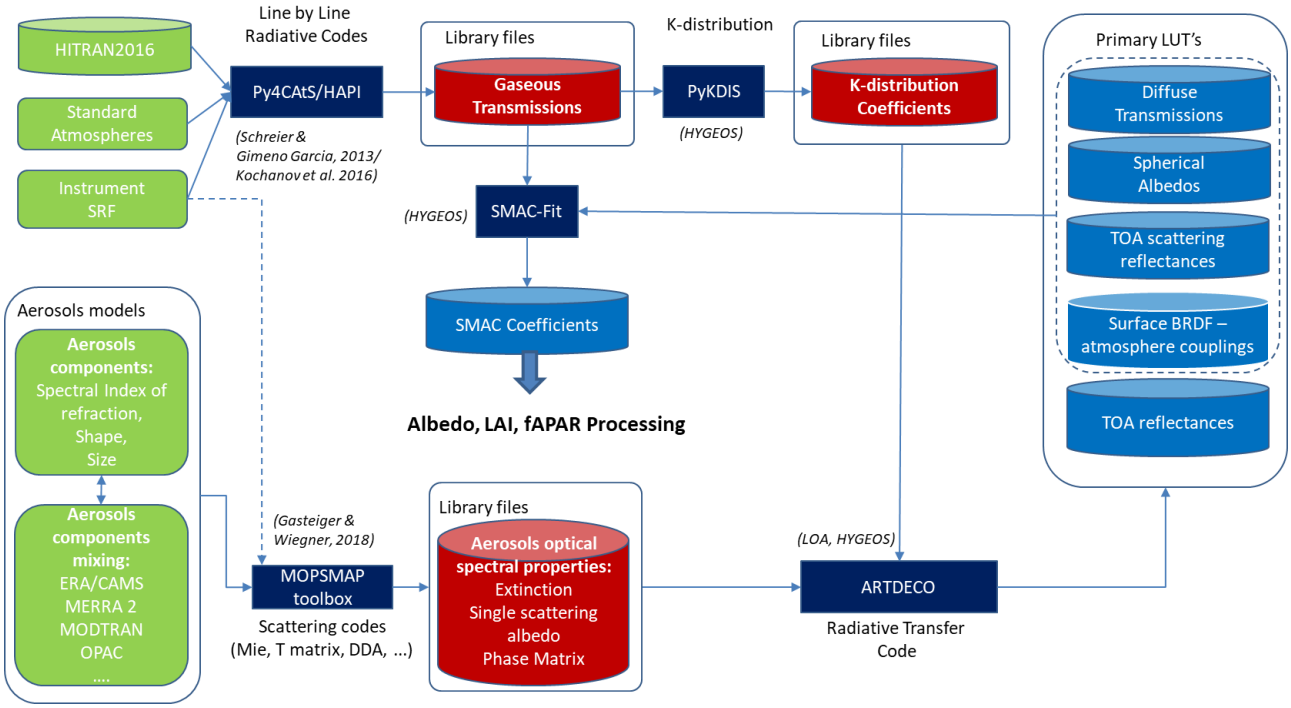

Figure 10 Organization of the radiative transfer tools at HYGEOS for computing atmospheric Look Up Tables from primary data

8.1.2 Gaseous absorption

With the SRF 's data as described in the previous section, we computed the typical absorption based on the spectroscopic HITRAN 2012/2016 database (Gordon et al., 2017) and the line-by-line radiative transfer tools Py4Cats (Schreier et al., 2013) and/or HAPI (Kochanov et al., 2016). Then regression of gaseous transmission versus gas concentrations, air masses and barometric pressure (see Figure 10) were computed to yield SMAC coefficients fitting the gaseous absorption. Correlated K-distribution coefficients computed with the new tool PyKdis, developed by HYGEOS, are also obtained in this step, for further use in the simulation of real TOA radiances for both sensors.

8.1.3 Aerosols optical properties

The optical properties for aerosols come from the OPAC (Hess et al., 1998) continental average model. This is a mixing of water soluble (N=7000), insoluble (N=0.4) and soot (N=8300) components. We created the optical properties (Mie computation) for each component separately and at each wavelength using the following refractive index computed for a relative humidity of 80%. This fix value is a starting point. It is planned to compute aerosols coefficients for Organic Matter, Sulphates Aerosol, Sea Salt and Desert aerosols in a future version.

8.1.4 Radiative Transfer computations

Radiative transfer equation for each pixel (geometry, band, AOT, Psurf) is solved independently using ARTDECO RT package (http://www.icare.univ-lille1.fr/projects/artdeco). The TOA reflectance computations are done for a black surface and no gaseous absorption to simulate scattering quantities used in 6S/SMAC, i.e.:

- TOA reflectance (ρ^{toa}) for aerosols and Rayleigh mixing, Aerosols and Rayleigh only
- Total (direct + diffuse) Transmission
- Spherical albedo

For each band, those optical quantities are simulated for various geometries, surface pressure, aerosol optical thicknesses… For example, for the TOA reflectance:

- RAA: 121 values from 0 to 180°
- SZA : 38 values from 0 to 89° (but used only if <70°)
- VZA: 47 values from 0 to 89° (but used only if <70°)
- Surface pressure: 6 values from 600 to 1050 hPa
- AOT at 550 nm: 16 values from 0.01 to 1.5 (but used only if < 0.8)

8.2 Annex 2: Aerosols models

Based on a statistical analysis of the occurrence of the 5 aerosols components (DUst, Black Carbon, Organic Carbon, Sea Salt, SUlfate) optical depths at 550 nm as reanalyzed by MERRA-2 globally, we built 148 mixtures with relative fractions X_u [u derived from the relative AOD at 550 nm:

X_u =

Then another assumption is done to compute the optical properties of the different mixtures. Each component is assigned an OPAC aerosol model as follows:

SU <-> OPAC antarctic aerosol model (ANTARCTIC)
DU <-> OPAC desert aerosol model (DESERT)
OC <-> OPAC water soluble component (WASO)
SS <-> OPAC maritime clean aerosol model (MARITIME)
BC <-> OPAC soot component (SOOT)

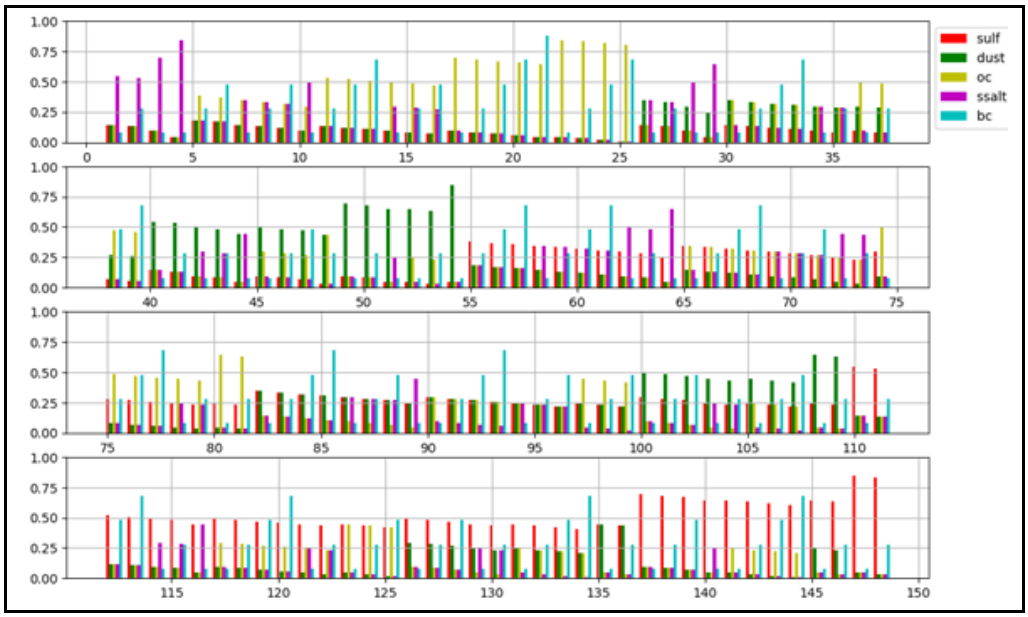

Figure 11: Fraction of the 5 MERRA-2 aerosols basic components of the 148 different aerosol mixtures used in the algorithm. (BC fraction is multiplied by 4 on this plot to be seen).

On Figure 12 is shown the aerosol basic optical properties of the 5 OPAC aerosol models used here for a set of Proba-V wavebands. On Figure 6 is shown the same optical properties but for the 148 aerosol mixtures.

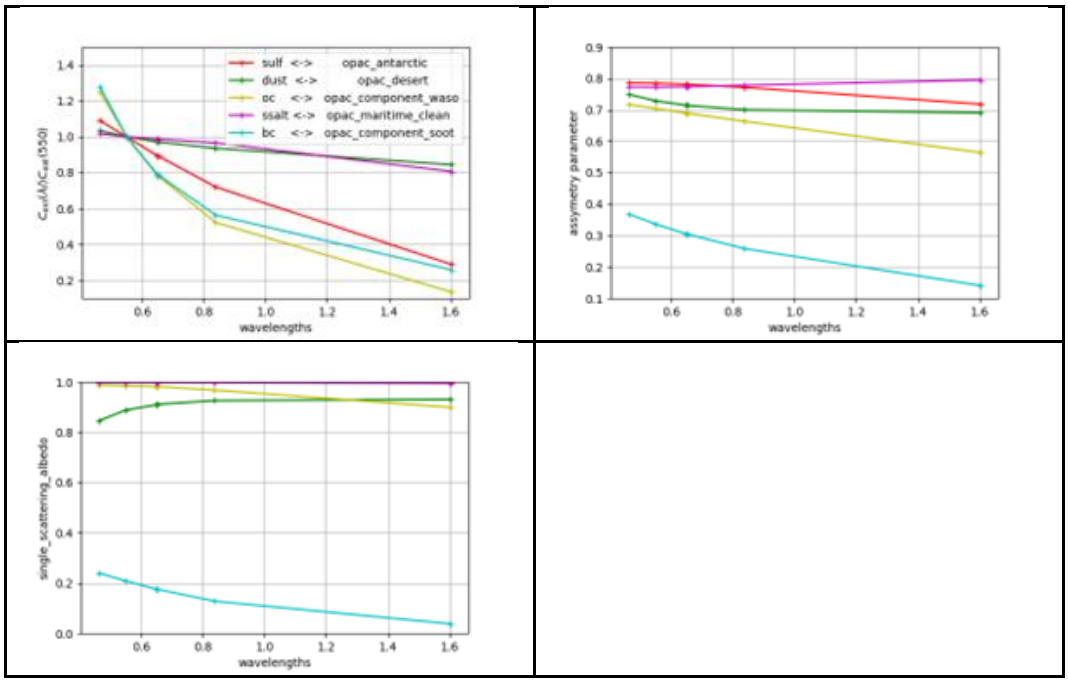

Figure 12: Optical properties of the 5 OPAC aerosol models for the Proba-V wavebands: (i) Extinction coefficient spectral dependency, (ii) asymmetry parameter and (iii) single scattering albedo.

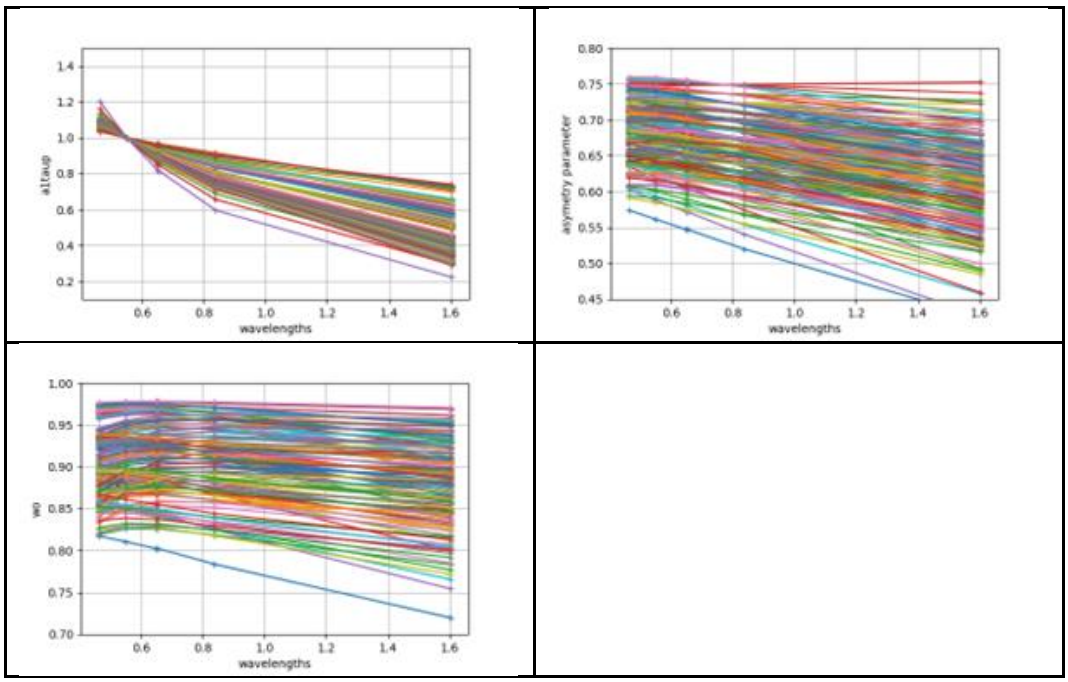

Figure 13: Same as Figure 12 but for the 148 different aerosol mixtures.