PROBA-V COLLECTION 1 USER MANUAL

PRODUCTS USER MANUAL

Reference: ***PROBA-V Products User Manual v3.03***
Author(s): Erwin Wolters, Wouter Dierckx, Marian-Daniel Iordache, and Else Swinnen
Version: 3.03
Date: 27/06/2023

DOCUMENT CONTROL

Signatures

Authors	Erwin Wolters, Wouter Dierckx, Marian-Daniel Iordache, and Else Swinnen
Reviewers	Carolien Toté, Dennis Clarijs
Approvers	Dennis Clarijs, Fabrizio Niro (ESA), Roberto Biasutti (ESA)
Issuing authority	VITO, authorized by ESA

Change record

Release	Date	Updates	Approved by
1.0	14/08/2014	Initial external version	Bart Deronde
1.1	07/10/2014	Added text on reprocessing, information on new PDP release	Bart Deronde
1.2	11/03/2015	Additions to PDP release, cloud detection, data support	Bart Deronde, Roberto Biasutti
1.3	31/08/2015	Matlab reader Update on product portal, mission extension, GeoTiff format, 1 km and 300 m projection	Bart Deronde, Roberto Biasutti
2.0	30/09/2016	Updated geolocation accuracy values Updated SRF figures and additional SRF figures per camera in Appendices Description of Level 2A algorithm, data, and metadata Description of modified cloud detection algorithm, with additional details included in Appendices Note on limited SZA range for atmospheric correction algorithm Description of reprocessing campaign Note on more convenient user registration for 1 km data Mapping information and table for GeoTiff files NDVI product description and scaling information Included Table with scale, offset, and no data values for all dataset types Description on inclusion of CF compliant metadata Figure update on opening HDF5 files in Quantum GIS	Dennis Clarijs, Roberto Biasutti

2.1	6/2/2017	Update Figure 4 with Level 2A data as end product Updated Figure 14: improved image resolution Included subsection on PROBA-V Mission Exploitation Platform (Section 3.2) Additional information on Level 2A GeoTiff files (Section 4.2). Various minor editorial changes.	Dennis Clarijs, Rosario Quirino Iannone
2.2	3/8/2017	Updated LTDN plot until March 2023 + explanation (Figure 1) Updated geolocation accuracy values for 16 June 2016 – 15 June 2017 (Table 2) Additional information on the PROBA-V MEP: pixel support (Section 3.2) Added short description and link to SNAP PROBA-V Toolbox (Section 4.6.10) Explanation of NDVI outliers (NDVI < -0.08 and NDVI > 0.92, Section 4.4.3)) Included R and Python code examples on how to read the Status Map (Section 4.6) Added short introduction on Quality Webpage (Section 5.3) Editorial corrections on metadata tables in Appendix D and added reference to CF compliant metadata introduction in Section 4.5.	Dennis Clarijs, Fabrizio Niro, Roberto Biasutti
3.0	21/2/2018	Updated List of Acronyms Updated geometrical accuracy values in Table 2 (Section 1.2, p. 17) Moved Collection 0 cloud detection algorithm description to Appendix D. Added disclaimer on difference in cloud cover percentage as displayed on Product Distribution Portal and in HDF5 metadata (Section 3.1.3, p. 37) Added Table with PROBA-V products spatial coverage on northern hemisphere	Dennis Clarijs, Fabrizio Niro, Roberto Biasutti

		winter and summer solstice (Section 2.4, p. 31) Replaced reference of Toté et al. (2016a) with journal paper of Toté et al. (2017, Section 2.5, p. 31) Updated MEP section with information on the Notebooks application (Section 3.2.5, p. 42). Added notification on V102 files (Section 4.3, p.46) Corrected filename convention for Level 1C filenames (Section 4.4.1, p. 49) Added link to ENVI documentation (Section 4.6.2, p. 52) Several editorial and lay-out corrections.	
3.01	16/3/2018	Included text on the availability of Antarctica data	Dennis Clarijs
3.02	27/3/2023	Added note on Proba-V Collection 2	Dennis Clarijs
3.03	27/6/2023	Added note related to Proba-V Collection 1 products no longer available to users.	Dennis Clarijs

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

Acronym	Explanation
API	Application Programming Interface
AU	Astronomical Unit
BOA	Bottom-of-Atmosphere
CCI	Climate Change Initiative
CESBIO	Centre d'Études Spatiales de la Biosphère
CF	Climate and Forecast
CGLS	Copernicus Global Land Service
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station
CRS	Coordinate Reference System
DEM	Digital Elevation Model
DMP	Dry Matter Productivity
DN	Digital Number Count
ECMWF	European Centre For Mid-Range Weather Forecasts
ENVI	Environment for Visualizing Images
EOS	Earth Observing System
ESA	European Space Agency
FAPAR	Fraction of Absorbed Photosynthetically Active Radiation
FCOVER	Fraction of green Vegetation Cover
FTP	File Transfer Protocol
FWHM	Full Width at Half Maximum
GAUL	Global Administrative Unit Layer
GDAL	Geospatial Data Abstraction Layer
GeoTiff	Geospatial Tagged Image File Format
GIS	Geographic Information System
GLCF	Global Land Cover Facility
GLSDEM	Global Land Survey Digital Elevation Model
HDF	Hierarchical Data Format
HDFS	Hadoop Distributed File System
HTTP	HyperText Transfer Protocol
ICP	Instrument Calibration Parameters
IDL	Interactive Data Language
IGFOV	Instantaneous Geometric Field Of View
LAI	Leaf Area Index
LEVEL 1C	Radiometrically and geometrically calibrated Level-1 data
Lref	Top-of-Atmosphere Reference Irradiance
LSB	Least Significant Bit
LTDN	Local Time of Descending Node
MATLAB	MATrix LABoratory
MEP	Mission Exploitation Platform
MERIS	Medium Resolution Imaging Spectrometer
MSB	Most Significant Bit
MVC	Maximum Value Composite
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NetCDF	Network Common Data Form
NFS	Network File System

NIR	Near-Infrared
NWP	Numerical Weather Prediction
OGC	Open Geospatial Consortium
PDP	Product Distribution Portal
PPT	PROBA-V Product Customization Tool
PROBA-V	Project for On-Board Autonomy - Vegetation
PV	Physical Value
QGIS	Quantum GIS
RoI	Region of Interest
RSS	Research and Service Support
S1	1-day Synthesis Products
S10	10-day Synthesis Products
SAD	Spectral Angular Distance
SMAC	Simplified Model for Atmospheric Correction
SNAP	Sentinel Application Platform
SNR	Signal-To-Noise Ratio
SPIRITS	Software for the Processing and Interpretation of Remotely sensed Image Time Series
SPOT-VGT	Satellite Pour l'Observation de la Terre – Végétation
SRF	Spectral Response Function
SSH	Secure Shell
SWIR	Short-Wave Infrared
SZA	Solar Zenith Angle
TOA	Top-Of-Atmosphere
TOC	Top-Of-Canopy
TOMS	Total Ozone Monitoring Spectrometer
USGS	United States Geological Survey
UTC	Universal Time Coordinate
VM	Virtual Machine
VNIR	Visible and Near-InfraRed
VZA	Viewing Zenith Angle
WGS84	World Geodetic System 1984
WKT	Well-Known Text
WMS	Web Mapping Service
WMTS	Web Map Tile Service
WRS-2	Worldwide Reference System V2

OBJECTIVES AND REFERENCE DOCUMENTATION

This document describes the PROBA-V product chain, the derived products, and the Product Distribution Portal at which the products are disseminated. The objectives of this document are the following:

- To present an overview of the PROBA-V satellite constellation and the measurement principles
- To provide an overview of the processing chain of the various PROBA-V products
- To give a detailed overview of the various datasets and product file attributes
- To guide the user through the registration and data ordering process
- To guide the user in the data viewing and handling

We have attempted to keep the document concise and comprehensible. Interested users on the various PROBA-V topics highlighted in this document are referred to the following scientific publications; see the References section for their full citations.

Document name	Major topics covered	Download location
Dierckx, W. et al. (2014). PROBA-V mission for global vegetation monitoring: standard products and image quality. *Int. J. Remote Sens*, **35**, 2589 – 2614.	PROBA-V mission, data quality, data compression, cloud detection, spectral response in relation to SPOT-VGT	http://proba-v.vgt.vito.be/sites/default/files/dierckx_etal_2014.pdf
Sterckx, S., et al. (2014). The PROBA-V mission: image processing and calibration. *Int. J. Remote Sens.*, *35*(7), 2565 – 2588.	PROBA-V mission, detailed processing chain overview, radiometric and geometric calibration, product distribution	http://proba-v.vgt.vito.be/sites/default/files/sterckx_etal_2014.pdf
Francois, M., et al. (2014). The PROBA-V mission: The space segment. *Int. J. Remote Sensing*, **35**, 2548 – 2564, doi:10.1080/01431161.2014.883098.	PROBA-V flight segment, instrument design, technology payloads, geometry and radiometry	http://proba-v.vgt.vito.be/sites/default/files/francois_etal_2014.pdf

1. Introduction

1.1. PROBA-V mission overview

The PROBA-V satellite was launched on 6 May 2013 and was designed to bridge the gap in space-borne vegetation measurements between SPOT-VGT (March 1998 – May 2014) and the Sentinel-3 satellites, of which the first is in orbit since 16 February 2016 and the second is planned for launch on 30 March 2018. The PROBA-V mission objective is to ensure continuity with the SPOT-VGT mission's heritage. The PROBA-V mission had a designed life time of 2.5 years, but the platform performance (LTDN evolution, payload performance, etc.) was well within requirements and in May 2015 it was decided to extend the mission with another 2.5 years until May 2018. Due to further excellent instrument and platform performance, another mission extension through October 2019 was decided on in May 2017.

The VEGETATION instrument onboard PROBA-V has a volume of just over 0.05 m^3 and weighs only 33 kg. PROBA-V flies at an altitude of 820 km in a sun-synchronous orbit with a local overpass time at launch of 10:45 h. After launch, the local overpass time first increased to 10:50 h in October 2014, followed by a decrease to 10:45 h in June 2016. Because the satellite has no onboard propellant, the overpass time will continue to decrease as a result of increasing atmospheric drag. Figure 1 presents the predicted Local Time of Descending Node (LTDN) evolution from May 2016 through March 2023. The horizontal and vertical lines show at which date the LTDN will be at a certain threshold value. By End-of-Mission in October 2019, the LTDN will be at ~09:45 h.

The VEGETATION instrument has a Field Of View of 102°, resulting in a swath width of 2295 km. This swath width ensures a daily near-global coverage (90%), whereas the full global coverage is achieved every 2 days. The central camera observes at 100 m nominal resolution, which covers a swath of about 517 km that ensures global coverage every 5 days.

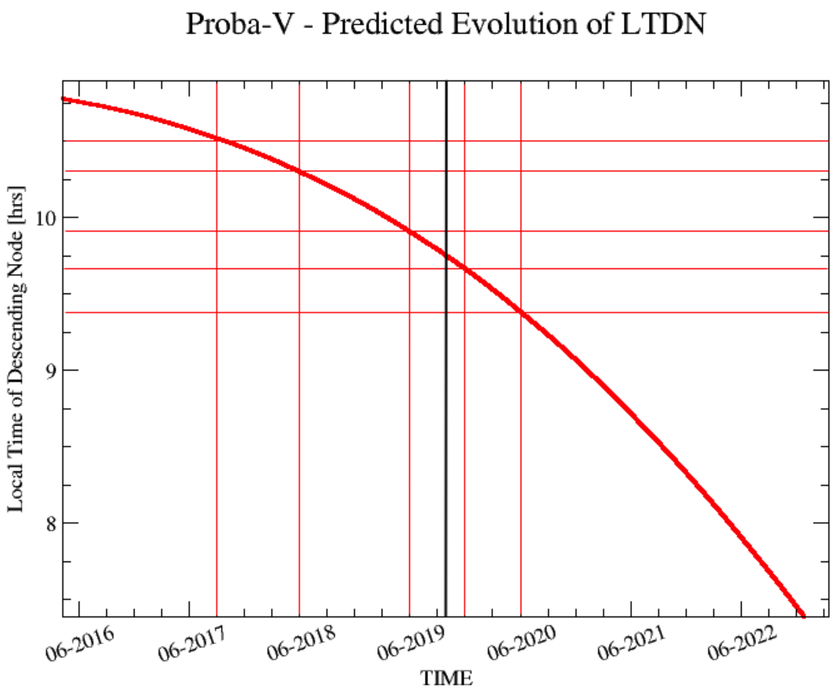


Figure 1: Predicted evolution of the PROBA-V local overpass time (Local Time of Descending Node, LTDN) from May 2016 through March 2023. Horizontal and vertical lines show the intersection of a certain LTDN with date. The thick black line indicates the October 2019 End-of-Mission date.

PROBA-V observes in four spectral bands: BLUE (centered at 0.463 µm), RED (0.655 µm), NIR (0.837 µm), and SWIR (1.603 µm). Observations are taken at resolutions between 100 and 180 m at nadir and up to 350 m and 660 m at the swath extremes for the VNIR and SWIR channels, respectively (Francois et al., 2014). Final PROBA-V products are disseminated at 100 m, 300 m and, 1 km resolution. The instrument and spectral characteristics will be explained in more detail in Section 1.2. The flight and payload characteristics are summarized in Table 1.

Table 1: PROBA-V payload and flight characteristics.

Altitude [km]	819 – 827
Local overpass time at launch [h]	10:45
Inclination [º]	98.7
Daily coverage [%]	90 (100 for latitudes > ±35º)
Payload Mass [kg]	33.3
Payload Dimensions [m]	0.2 × 0.8 × 0.35
Designed lifetime [yr]	2.5 – 5
Instantaneous geometric field of view (IGFOV) [m]	96.9 for VNIR (BLUE, RED, NIR), 193.8 for SWIR

1.2. Instrument characteristics

The optical design of PROBA-V consists of three cameras. Each camera has two focal planes, one for the short wave infrared (SWIR) and one for the visible and near-infrared (VNIR) bands. The VNIR detector consists of four lines of 5200 pixels. Three spectral bands were implemented, comparable with SPOT-VGT: BLUE, RED, and NIR. The SWIR detector is a linear array composed of three staggered detectors of 1024 pixels. Each used detector line is labelled as a strip. Each camera therefore has 6 strips. The instrument plane layout is shown in Figure 2.

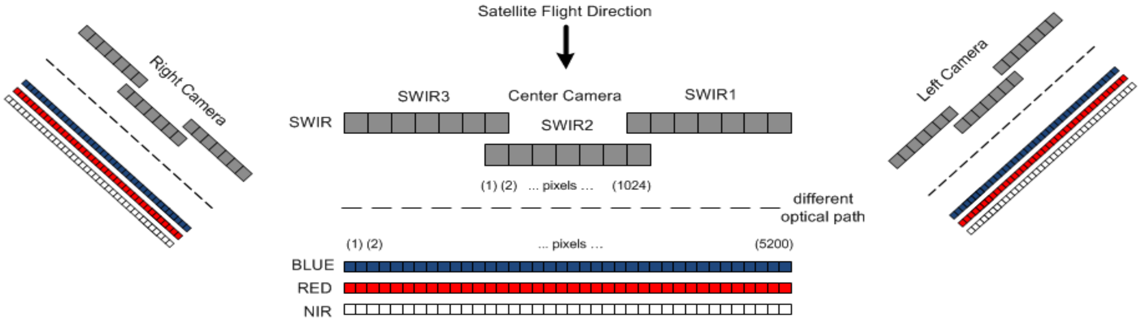


Figure 2: PROBA-V instrument layout.

The instrument has been designed such that the NIR band observes the Earth first, followed by the RED, BLUE, and SWIR bands. As a result, an observation time difference of 12 s exists between the NIR and SWIR bands. This difference is accounted for in ground surface observations, however, it impacts the cloud detection, which will be further discussed in Section 2.2.3. Table 2 lists the radiometric characteristics of the PROBA-V spectral bands.

Table 2: PROBA-V spectral, radiometric, and geometric characteristics. L_{ref} refers to the Top-Of-Atmosphere (TOA) irradiance at the respective spectral band. Geometric mean accuracy values were obtained over the period 16 December 2016 – 15 December 2017. FWHM = Full Width at Half Maximum, SNR = Signal to Noise Ratio.

Band name	**Centre wavelength [µm]**	**Spectral range @FWHM [µm]**	**SNR @L_{ref} [W m^{-2}sr^{-1} µm^{-1}] at 300 m resolution**
BLUE	0.464	0.440 – 0.487	177 @111
RED	0.655	0.614 – 0.696	598 @110
NIR	0.837	0.772 – 0.902	574 @106
SWIR	1.603	1.570 – 1.635	720 @20
Radiometric performance			
Absolute accuracy [%]	< 5		
Inter-channel accuracy [%]	< 3		
Stability [%]	< 3		

Geometric performance	
Mean geolocation accuracy (standard deviation) [m]	BLUE:77.6 (92.6) RED: 73.2 (79.0) NIR: 69.2 (77.9) SWIR: 71.4 (78.7)

Figure 3 presents the spectral response functions (SRFs) for the PROBA-V BLUE, RED, NIR, and SWIR channels (solid lines), SPOT-VGT1 (dashed lines), and SPOT-VGT2 (dotted lines). It can be seen that differences between the PROBA-V and SPOT-VGT SRFs exist and that these differences are largest for the SWIR band. Note that the spectral responses for PROBA-V represent the center camera and that slight differences between the left, center, and right cameras exist. Appendix A shows detailed plots with spectral responses for all PROBA-V cameras.

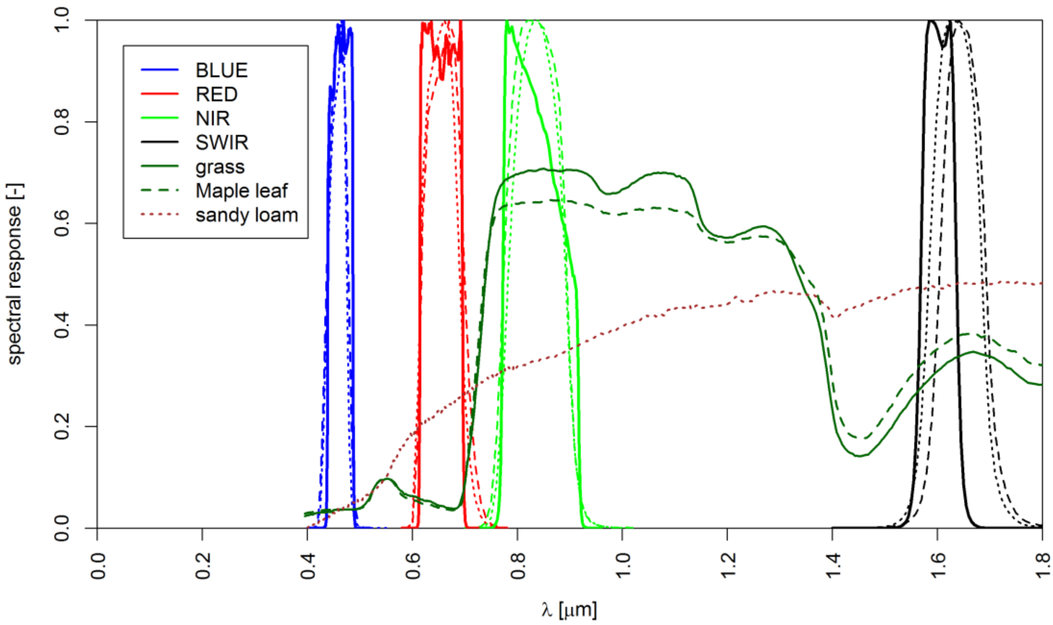


Figure 3: Spectral response functions for SPOT-VGT1 (dashed lines), SPOT-VGT2 (dotted lines), and PROBA-V (solid lines) for the BLUE, RED, NIR, and SWIR channels. Typical vegetation spectra for grass (solid dark green line), maple leaf (dashed dark green line), and bare soil (sandy loam, dotted brown line) are plotted for reference.

1.3. PROBA-V data products

The PROBA-V products are similar to the ones of SPOT-VGT in terms of file structure and comprise the following elements:

- **Segment products (Level 1C and Level 2A, both consisting of TOA reflectances)**
  The Level 1C product contains the raw, unprojected observations in segments, as well as calibration information, while the Level 2A (L2A) products contain the projected segment data. These latter data were named “P-products” in the SPOT-VGT era.

- **Synthesis products (Level 3, both TOA and TOC)**
  These products contain daily (S1, available at all resolutions) and multi-daily (S5 for 100 m and S10 for 300 m and 1 km) TOA reflectances that are composed of cloud, shadow, and snow/ice screened observations. Additionally, Top-of-Canopy (TOC) reflectance and NDVI products are corrected for atmospheric reflectance contributions, such as aerosols and gaseous absorption. Synthesis products were previously known as S-products for SPOT-VGT.

Figure 4 shows the flowchart of the product processing chain. The separate products and algorithms will be further described in Section 2.

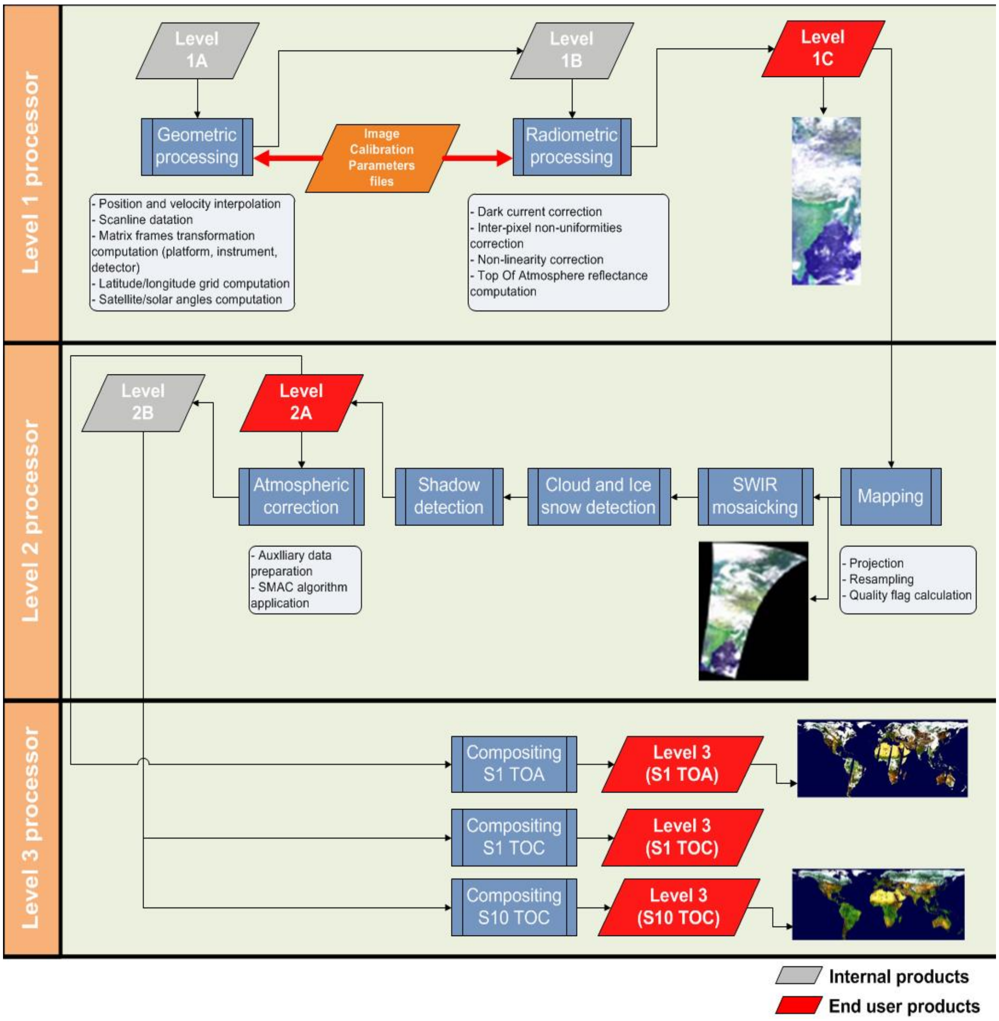


Figure 4: PROBA-V processing chain flowchart.

This Product User Manual concerns the PROBA-V Collection 1 products. Since March 2023, PROBA-V Collection 2 is available, with improved products after full reprocessing of the PROBA-V Collection 1 archive. For more information, see https://proba-v.vgt.vito.be/en.

2. Products description

This Section describes the various PROBA-V products. First, the various algorithms that are applied to the raw image data are explained, followed by an explanation of the compositing rules to arrive at the Level 3 synthesis products. Finally, for all product types an overview of the information content is given.

2.1. Level 1 algorithm and data

The upper part of Figure 4 (*'Level 1 processor'*) shows the subsequent processing steps, which are performed to obtain the Level-1C product. The two main processing steps are:

- Geometric processing
- Radiometric processing

These processing steps are explained in further detail in the following subsections.

2.1.1. Geometric processing

Using the Level 1A raw and uncompressed data, a geolocation step is performed for each satellite position to determine the latitude and longitude of the observed pixel. The satellite position and velocity are interpolated for each scan line using an orbital propagation model. The geolocation accuracy is refined using the geometric Instrument Calibration Parameters (ICP) file (see also Figure 4). The ICP file contains the variation in detector viewing direction relative to the time out of eclipse and the Sun beta angle. The geometric processing model additionally calculates the viewing and solar zenith angles (VZA and SZA, respectively), which are required for further processing. The output of the geometric processing are the Level-1B data. The user is referred to Sterckx et al. (2014) for further details on the geometric processing model.

2.1.2. Radiometric processing

The radiometric processing converts the digital number count at a certain spectral band (DN) into physical TOA reflectance values. First, the DN number is corrected for detector non-linearities, dark currents, and inter-pixel non-uniformities. Second, these numbers are converted to at-sensor radiance L [W m^{-2} µm^{-1} sr^{-1}], using the band-specific calibration coefficients derived from the radiometric ICP file. Finally, the TOA radiance L at a given spectral band is converted into TOA band reflectance using:

$$R_{TOA} = \frac{\pi \times d^2 \times L}{E_o \times cos(\theta_s)}$$

With R_{TOA} the obtained TOA reflectance value [-], *d* the Earth – Sun distance [AU], E_o the mean exo-atmospheric irradiance at the specific spectral band [W m^{-2} µm^{-1}], with values from Thuillier et al. (2003), and θ_s the solar zenith angle [°]. The output of the radiometric processing are the Level 1C data.

2.2. Level 2 algorithm and data

The Level 1C data are used as input for further processing in the Level 2 processor, which consists of the following steps:

- Mapping and SWIR mosaicking
- Snow/ice detection
- Cloud and cloud shadow detection
- Atmospheric correction

Please note that the compositing procedure for the 300 m and 1 km products differ in certain steps, which will be explained in more detail in Section 2.5. The separate processing steps are explained in the following subsections.

2.2.1. Mapping and SWIR mosaicking

In the mapping procedure, the Level 1C data are mapped onto a World Geodetic System (WGS) 84 geographic lat/lon projection, using a procedure proposed by Riazanoff (2004). An inverse model is used to calculate per pixel the original Level-1 *(p, l)* coordinates from the Level-2 *(x, y)* coordinates, with *x* being the longitude, *y* the latitude, *p* the pixel-in-line, and *l* the line number. This mapping is explained in
Figure 5.

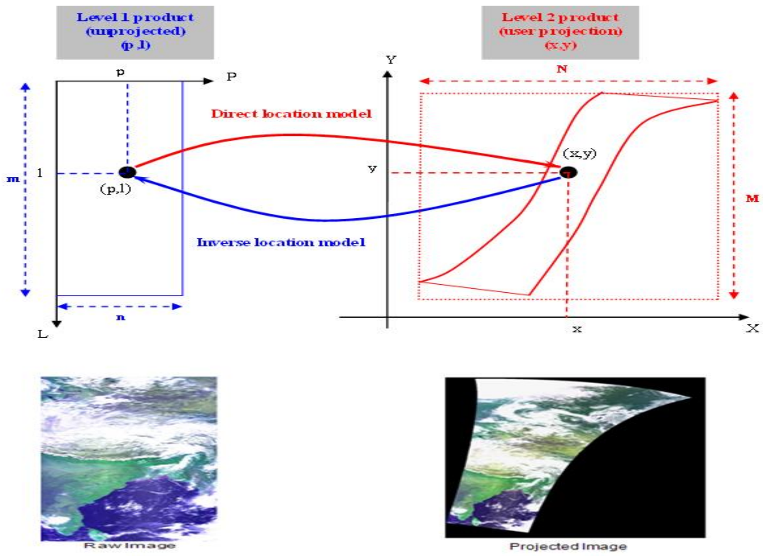


Figure 5: The Level 2 mapping procedure.

The mapping operation is carried out twice, at 0 m and 5000 m above sea level, thereby resulting in two *(p, l)* coordinate sets. The *(p, l)* coordinates at a given altitude are then linearly interpolated from these two datasets. Ortho-rectification is performed using the Global Land Survey Digital Elevation Model from the National Aeronautics and Space Administration (NASA)/United States Geological Survey (USGS) Digital Elevation Model (GLSDEM). More information on the GLSDEM can be found at http://glcf.umd.edu/data/glsdem/ and data can be freely downloaded from the Global Land Cover Facility (GLCF) FTP site: ftp://ftp.glcf.umd.edu/glcf/GLSDEM/. The data have a resolution of ~90 m and are available in Worldwide Reference System version 2 (WRS-2) format or in degree tiles for the latitudinal range 56°S – 83°N.

In the final step, the Level 2 pixel values are mapped to an *(x,y)* grid using a stretched bi-cubic interpolation filter (see Dierckx et al., 2014). This interpolation technique was found to be more accurate for PROBA-V compared to the standard bi-cubic interpolation used for SPOT-VGT1 and SPOT-VGT2 (Dierckx et al., 2014). The SWIR detector per camera consists of three strips (see Figure 2). After the mapping, there are still three separately projected SWIR strips. Therefore a mosaicking step is applied to compose a single SWIR band image. In the overlapping regions the pixel radiometric Status Map is taken into account to select the best pixel (see
Figure 6). More information on the Status Map dataset is given in Section 5.

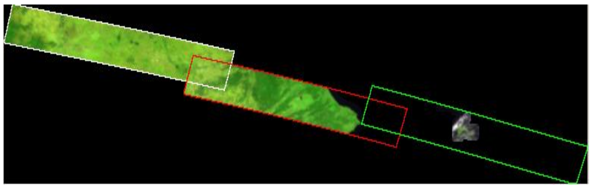


Figure 6: Example of the mosaicking algorithm result on the three SWIR strips.

2.2.2. Snow/ice detection

The snow/ice detection for PROBA-V is similar to the approach in the SPOT-VGT mission. The binary snow mask uses five indices based on the TOA reflectance observed in the four PROBA-V spectral bands:

$$T_1 = R_{RED}$$

$$T_2 = R_{SWIR}$$

$$T_3 = \frac{R_{BLUE} - R_{NIR}}{R_{BLUE} + R_{NIR}}$$

$$T_4 = \frac{R_{BLUE} - R_{SWIR}}{R_{BLUE} + R_{SWIR}}$$

$$T_5 = \frac{R_{BLUE} + R_{RED}}{2} - R_{SWIR}$$

Figure 7 shows the decision tree, which maps the input to the final classification by performing a sequence of checks on the different indices values.

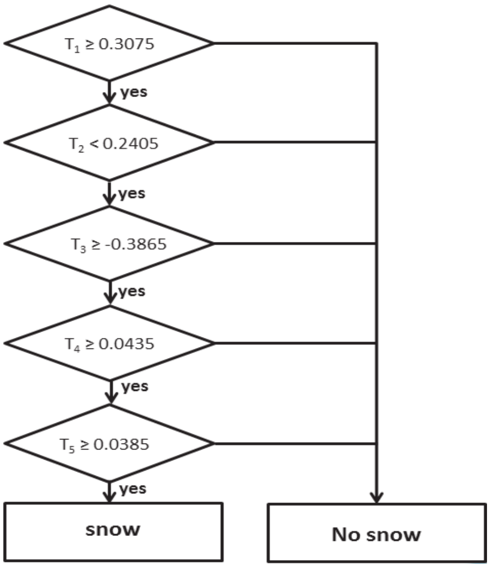


Figure 7: Snow/ice detection decision tree.

2.2.3. Cloud and cloud shadow detection

2.2.3.1. Cloud detection

Clouds obstruct land surface parameter retrieval in satellite observations. Therefore a proper cloud screening is pivotal in the pre-processing for the various value-added products.

Many studies, as well as user feedback identified several issues with the Collection 0 PROBA-V cloud detection algorithm. The Collection 0 algorithm is based on the use of static thresholds applied to the BLUE and SWIR spectral bands. False cloud detection over bright surfaces, such as deserts and salt lakes, and flagging of thick ice clouds as 'snow/ice' were among the key problems of the operational cloud screening method. To overcome these limitations, a new algorithm was developed and implemented for the PROBA-V reprocessing (Collection 1).

The Collection 1 cloud detection algorithm is described below, readers interested in the Collection 0 cloud detection algorithm are referred to Appendix D. More information on the Collection 1 reprocessing campaign is given in Section 2.5.

Collection 1 cloud detection algorithm

The improved and currently operational cloud detection algorithm addresses the main limitations of the Collection 0 cloud detection algorithm by using a more extensive and sophisticated set of cloud tests.

The improved algorithm introduces major changes in the following aspects:

- A supervised training of a classification scheme that was designed to replace the operational Collection 0 algorithm.
- High-resolution surface albedo data are used as background reference maps.
- The decision to assign a pixel to 'cloud' or 'clear' is made via an extended set of threshold tests and similarity checks.

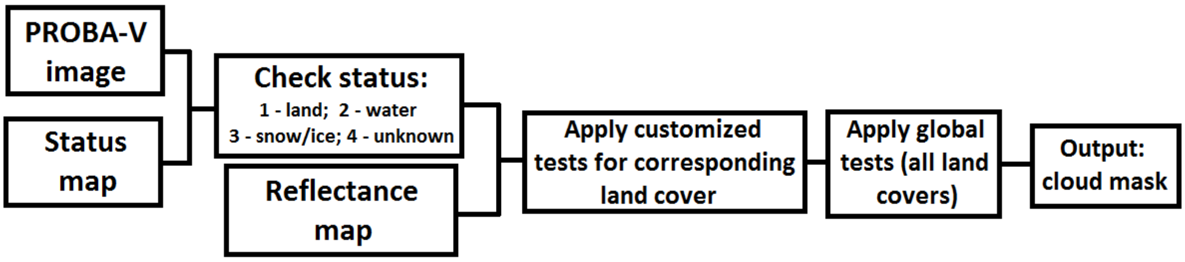


Figure 8: Flowchart of the Collection 1 cloud detection algorithm.

A flowchart of the Collection 1 cloud detection algorithm is presented in Figure 8. First, each PROBA-V pixel is assigned to a land cover class based on monthly images generated by ESA's Land Cover Climate Change Initiative[1] (CCI), of which the classes 'land', 'water', 'snow/ice', and 'unknown' are used for further processing. Subsequently, for each land cover class, background surface reflectances for the BLUE spectral band are generated, based on a monthly clear-sky climatology obtained from Medium Resolution Imaging Spectrometer (MERIS) 0.413 and 0.443 µm observations over the period 2002 – 2012. In case of missing data (e.g. over areas in the winter season), coarse-resolution (5 km) broad-band (0.3 – 0.7 µm) ESA's GlobAlbedo[2] surface reflectance data were used.

Additionally, reference spectra were built from clear-sky PROBA-V observations for specific land cover types. These spectra were built according to previous experience and literature reports on challenging cases, e.g. snow/ice areas, pixels with semi-transparent clouds, and salt planes. The concept of reference spectra, with the PROBA-V spectral bands superposed for convenience, is shown in Figure 9.

[1] European Space Agency, *CCI Land Cover Project – Algorithm Theoretical Baseline Document Version 2*, available online: http://www.esa-landcover-cci.org/?q=webfm_send/75
[2] European Space Agency (ESA), *GlobAlbedo Project,* http://www.globalbedo.org

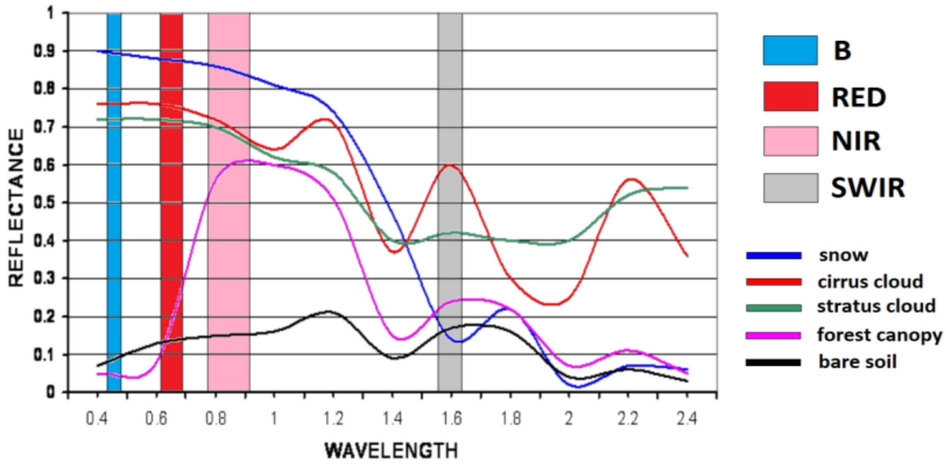


Figure 9: Representative reference land cover and cloud spectra with the PROBA-V spectral bands superposed. Figure adapted from Jedlovec, (2009).

From Figure 9, several generic observations can be made.

- clear pixels (except for snow/ice) have lower BLUE reflectances than cloudy or semi-transparent cloudy pixels
- the reflectance ratio BLUE / SWIR is smaller for clear pixels than for cloudy pixels (except for snow/ice)
- SWIR reflectances for cloudy pixels are larger than for clear pixels
- clear pixels (except for snow/ice) have a lower reflectance ratio BLUE / NIR than cloudy pixels;

In the final step, a set of decision rules were defined, consisting of threshold tests (on band reflectances, reflectance ratios or amplitude differences) and similarity checks. The similarity checks were performed using the Spectral Angular Difference (SAD), which measures the cosine of the angle between two vectors. Low SAD values indicate high similarity, while large SAD values show low similarity. The metric was chosen as it is ideally invariant with the illumination conditions. The following reflectances were assessed in the decision rules:

- the BLUE TOA pixel reflectance
- the SWIR TOA pixel reflectance
- the TOA pixel reflectance spectrum (based on the BLUE, RED, NIR, and SWIR reflectances)
- the BLUE TOC reflectance from the reference image, evaluated at the image pixel location using bilinear interpolation.

The decision rules (see Appendix B for more details) were tuned on a training dataset, which was randomly sampled from a seasonally and globally distributed pixel database. Initial validation on a set of manually cloud-screened PROBA-V images shows that the modified algorithm significantly improves on the current operational cloud detection, see Stelzer et al. (2016) for more details on the validation methodology and results.

During the final Collection 1 validation a cloud mask issue was found. This issue occurs every year during the winter period (December – January) at latitudes northward of ~50°N. The root cause of this issue was found in the presence of several pixels classified as “unknown” in the input land cover map used to drive the cloud detection algorithm, as well as through the usage of the GlobAlbedo surface albedo maps as fall-back in case of missing pixels in the background MERIS surface reflectance climatology. This results in cloud over-detection for pixels being classified as “unknown” and under-detection in case of a brighter GlobAlbedo background reflectance.

The cloud detection issue impacts both the segment data (Level 2A) and the synthesis data at all product resolutions and is most prominent during the northern hemisphere winter season.

2.2.3.2. Cloud shadow detection

Cloud shadow detection is also of importance to land surface research, as the dark areas casted at the Earth surface can lead to erroneous vegetation parameter retrievals. The methodology to screen for cloud shadows from PROBA-V observations is a hybrid between the radiometric approach (see e.g. Zhu and Woodcock, 2012 and Ackerman et al., 2010) and a geometric approach (see Simpson et al., 2000).

The geometric part of the cloud shadow detection algorithm is presented in Figure 10. A cloud pixel is located at position *p*, with the actual cloud being at height *h* from the tangential plane, i.e., the intersection of the sun beam and the line of sight from the satellite to the cloud pixel. The cloud shadow can then be found as the intersection of the sun beam and the tangential plane at the center. Solar zenith and azimuth angles are assumed to be equal in the cloud and cloud shadow pixels. It follows from Figure 10 that angle φ equals the sum of Υ and the viewing azimuth angle φ_{av}. When φ and the distance between the cloud and associated cloud shadow pixel, *r*, are known, their position can be calculated (see Sterckx et al., 2014 for further details).

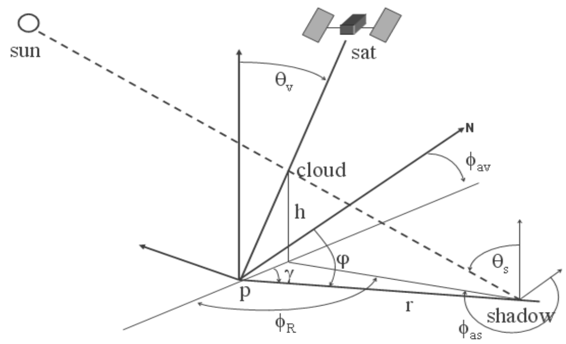


Figure 10: Depiction of solar, satellite, cloud, and cloud shadow geometries.

Cloud heights are estimated using the gradient in NIR reflectance along the projected path from a cloud to its shadow in the image (
Figure 11). In case of a cloud shadow, the NIR reflectance will decrease towards a minimum from cloud to shadow edge. If this change is above a threshold of 20%, a shadow edge is detected. From

the locations of the cloud and shadow edge, the cloud height can subsequently be calculated. More details on the cloud shadow detection can be found in Sterckx et al. (2014).

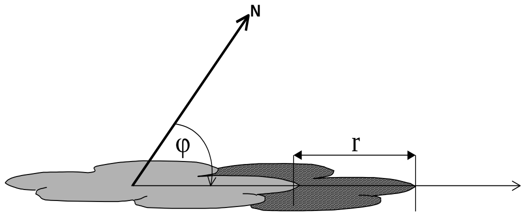


Figure 11: Concept of reflectance tracing along the cloud-to-shadow path.

2.2.4. Atmospheric correction

The Level 2A TOA reflectance observations are the resultant of surface reflectance and scattering, absorption, and multiple reflections within the atmospheric column below the satellite (clouds, gases, aerosols). In order to obtain the directional TOC reflectance values (Level 2B data), version 4.2 of the Simplified Model for Atmospheric Correction [SMAC, Rahman et al. (1994)] is used. This model converts the observed TOA reflectance into TOC reflectance using auxiliary water vapour, ozone, and surface pressure data. Water vapour content is taken from the European Center for Mid-Range Weather Forecasts (ECMWF) Numerical Weather Prediction (NWP) model delivered by MeteoServices (http://www.meteoservices.be), which is bi-linearly interpolated in space and linearly in time. For ozone, a climatology based on 11 years of Total Ozone Mapping Spectrometer (TOMS) observations prepared by the Centre d'Études Spatiales de la Biosphère (CESBIO) is used. Surface pressure is derived from the Global Land Surface Digital Elevation Model (GLSDEM), using a height to pressure conversion formula proposed by Plummer et al. (2003). The aerosol optical thickness (AOT) is retrieved using an empirical relation between TOA NDVI and the SWIR / BLUE TOC reflectance ratio. This aerosol retrieval can only be applied for pixels with sufficient vegetation (NDVI > 0.2 and TOC SWIR < 0.4), for pixels not fulfilling these criteria a simple AOT as function of the latitude is used (Berthelot et al., 1997):

$$AOT_{0.55} = 0.2(\cos(\varphi) - 0.25)\sin\,(\varphi + \frac{\pi}{2})^3 + 0.05$$

The SMAC algorithm uses a separate equation for each of the atmospheric interaction processes. Scattering and absorption by atmospheric constituents are parameterized by analytical formulations, whose coefficients are fitted against reference values derived by the 6S radiative transfer model (Vermote et al., 1997).

It is noted that, due to the limited validity range of the regression coefficients, the SMAC correction becomes less accurate, but still acceptable, for SZA or VZA > 30°, while the correction becomes unreliable for SZA or VZA > 60° (Proud et al., 2010). Therefore, TOC reflectances at such large angles (PROBA-V observes at SZA up to 87.3°) should be considered with caution. Observations with extremely high SZA (> 80°) are only included in the TOA data (Level 2A). See Table 4 (p.33) for information on the TOA and TOC products geographical extents.

2.3. Level 3 algorithm and data: compositing

The compositing into synthesis images is performed by the Level 3 Processor (see Figure 4). The aim is to optimally combine multiple observations into a single and cloud-free synthesis image. Atmospherically uncorrected (Level 2A) or corrected (Level 2B) data are the basis for the TOA and TOC synthesis products, respectively. Cloud coverage is minimized through discarding pixels that were labeled as cloudy by the cloud detection algorithm. In addition, angular variations are minimized, while global coverage is maximized. The S10 compositing is applied to avoid spatial coverage gaps resulting from clouds and the non-global daily swath coverage in the tropical areas.

Atmospherically corrected segment files are combined into a global Level-3 synthesis through application of a Maximum Value Composite (MVC) technique (see among others Holben, 1986 and Tarpley et al., 1984). This technique selects the maximum TOA NDVI (which is additionally calculated within the compositing algorithm) pixel values. The following two synthesis products are generated:

- S1 (1-day syntheses): TOA and TOC
- S10 (10-day or dekad syntheses): TOC, with starting days at the 1st, 11th, or 21st day of a month. For months having 28, 29 or 31 days the S10 of the third dekad comprises the remaining days of that month.
- For the 100 m product, also S5 TOA and TOC data files are available. PROBA-V 100 m S5 products are comparable with full-coverage 300 m S1 products and are not real syntheses. Due to the narrow swath of the 100 m camera, there is only overlap in observations for latitudes > ~40^{o}. This means that only poleward of this latitude compositing rules can be applied and that within ~40^{o} S – 40^{o} N the reflectances observed at one of the five days are given.

The TIME grid dataset in the S5 files provides information at which day observations over specific regions were performed. This information is provided in minutes since the start of the synthesis period (day 1, 00:00 UTC). Figure 12 indicates the 100 m observation coverage after 5 days. Please note that S5 data can only be ordered for day ($n \times 5$) + 1, with n=[1,5].

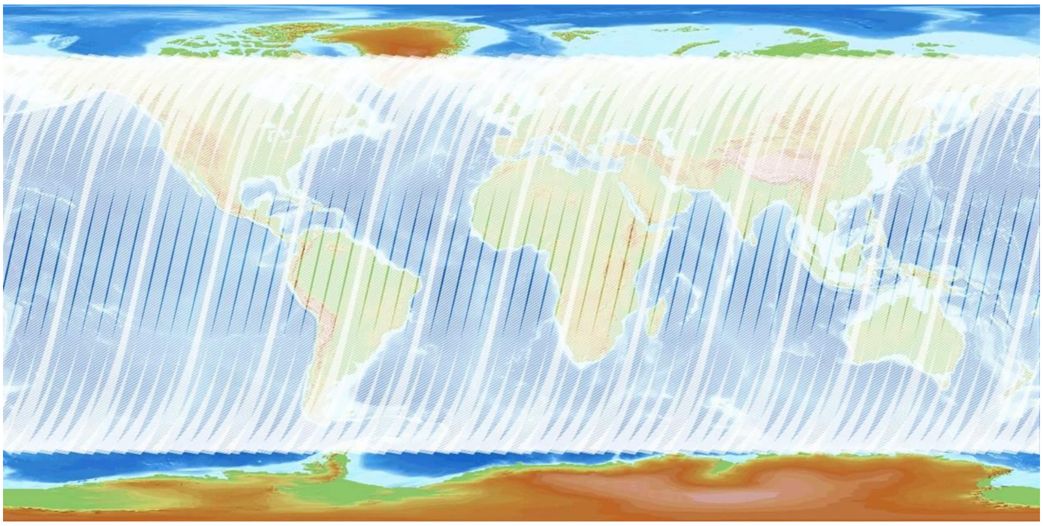


Figure 12: Overview of the 100 m coverage after 5 days. The brighter white areas indicate overlapping observations.

In order to preserve continuity between the PROBA-V and SPOT-VGT 1 km products, the compositing rules for the 1 km resolution differ from the 300 m resolution. For clarity's sake, the compositing rules for all resolutions are listed below.

The compositing rules **for the 300 m and 100 m** syntheses are as follows:

- Observations covered by all spectral bands are preferred over observations covered by only a few spectral bands.
- Observations with a good pixel quality indicator for all bands are preferred over observations of less quality.
- Cloud-free observations are preferred over ice/snow observations, which in turn are preferred over cloudy observations.
- In case two observations satisfy the rules above, the VZA and SZA are used to distinguish optimal from less optimal observations. The larger the VZA and/or SZA, the larger the (two-way) optical path length. Using the thresholds presented in Table 3, observations are categorized as 'good', 'acceptable', and 'bad'. Logically, the selection order is 'good' > 'acceptable' > 'bad' (See Figure 13 for the decision tree).
- In case two or more observations are still of equal quality, the observation yielding the maximum TOA NDVI value is preferred.

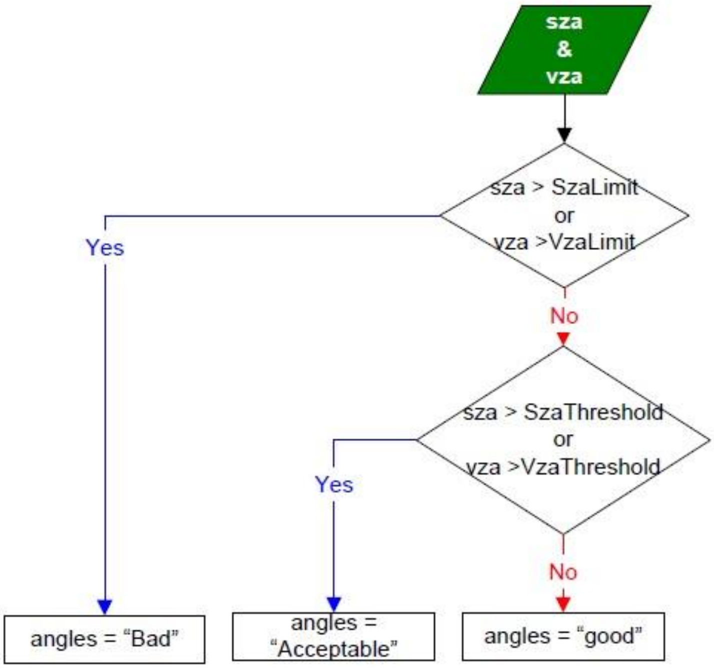


Figure 13: Procedure to determine the observation quality based on SZA and VZA in the synthesis processing. Note that in the 1 km processing none of these rules are applied.

Table 3: SZA and VZA threshold values in the synthesis processing. Note that in the 1 km processing none of these rules are applied.

Rule	Limit	Threshold
Solar Zenith Angle (SZA)	90°	60°
Viewing Zenith Angle (VZA)	75°	40°

The compositing rules **for the 1 km** syntheses are as follows:

- Observations covered by all spectral bands are preferred over observations covered by only a few spectral bands.
- Observations with a good pixel quality indicator for the BLUE, NIR, and RED bands are preferred over observations of less quality. <u>This differs from the 300 m compositing rule in that SWIR observations with lower than 'good' radiometric quality are allowed.</u>
- Cloud-free observations are preferred over ice/snow observations, which in turn are preferred over cloud observations.
- In case two or more observations are still of equal quality, the observation yielding the maximum TOA NDVI value is preferred.

It is noted that due to the compositing consistency with SPOT-VGT, neither the SZA nor the VZA selection rules are applied at 1 km resolution. As a result of these compositing rules, the 1 km synthesis products will sometimes contain pixels with a 'bad' SWIR status, while being cloud-free and having a 'clear and good' status for the other bands. These pixels can be identified in the synthesis status map by a status value of 232 instead of 248 (see Section 5.2).

Such pixels have been flagged because they have an unusually high dark current value compared to other SWIR pixels. In most cases, the pixel values involved are still reliable and are handled by the radiometric correction as part of the Level-1 processing. However, these pixels are considered by the PROBA-V Calibration Team to have a suboptimal pixel quality and should be treated as such by the user.

2.4. Data projection and geographical extents

All PROBA-V data products are projected in a standard WGS84 projection (also known as the Plate Carrée projection), similar as for the SPOT-VGT products. The 1 km Plate-Carrée projection is defined as 1/112°, with the latitude and longitude coordinates defined at the pixel centre. This implies that the pixel boundaries extend ± 1/224° for both latitude and longitude at the pixel corners. For example, if we consider the pixel corresponding to [lon, lat]=[-180°, 75°], the upper left corner of this pixel represents [lon, lat]=[-180° - 1/224°, 75° + 1/224°].

For the 300 m products, it seems logical to define a projection that contains 336 pixels per degree, such that 3 × 3 pixels would map onto a single 1 km pixel. However, users should note that due to the pixel coordinate definition (which applies to both 1 km and 300 m), no proper aggregation of 300 m to 1 km can be performed at the minimum and maximum latitude and longitude, while such an aggregation can be done within these boundaries (see the solid grids in Figure 14). Likewise, caution should be taken with the aggregation of 100 m pixels onto the 300 m grid.

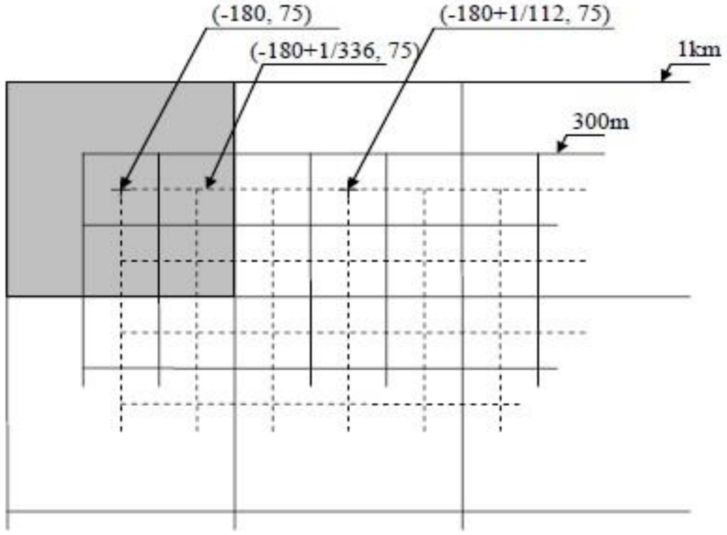


Figure 14: Depiction of the Plate-Carrée 1 km and 300 m projection grids. Solid lines indicate grids with coordinates representing the pixel centre, while for the dashed grid these represent the pixel upper-left corner.

Table 4 contains information on the geographical extents of the PROBA-V data products at the 21 June and 21 December solstices.

Table 4: Maximum northern latitude [º] for the PROBA-V data products at the summer and winter solstice dates.

Date	Level 2A	S1 TOA	S1 TOC	S5/S10 TOC
21 June	79.3	75	75	75
21 December	65.9	65.9	65.9	66.0/66.2

The larger geographical coverage for the Level 2A (TOA) products results from a decision to extend the observations to SZA = 87.3°N to provide additional observations for snow cover mapping. Figure 15 clearly shows the extended observations over the northern hemisphere for January – March 2015 (blue) and 2017 (orange). The extended observations are not included in the synthesis TOC products, due to the inaccurate atmospheric correction by SMAC at these large SZA (see Section 2.2.4).

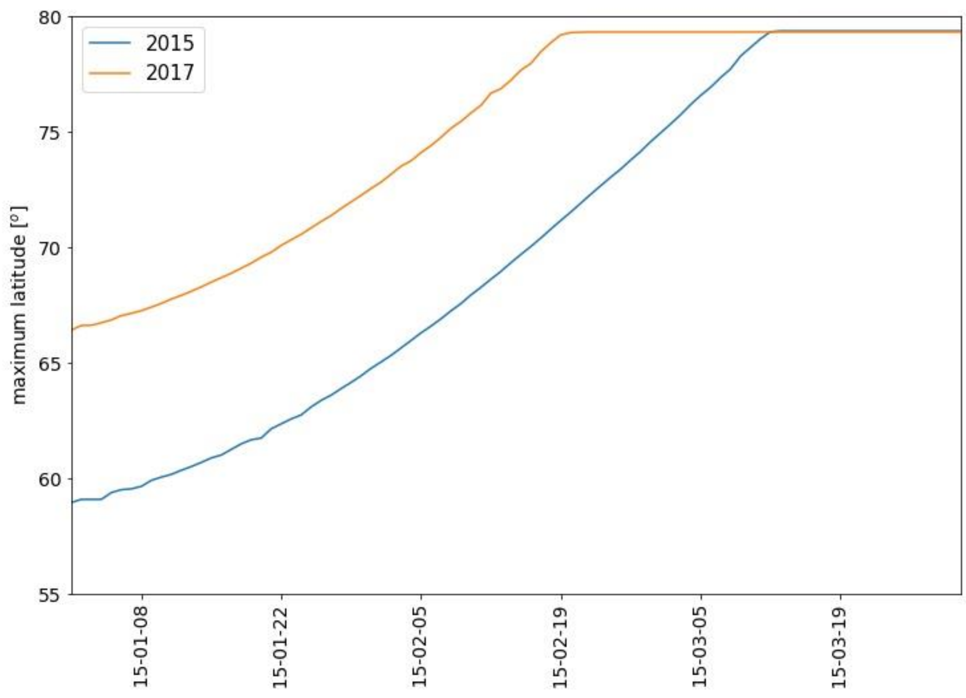


Figure 15: Northernmost Level 2A observations for January – March 2015 (blue) and 2017 (orange).

2.5. 2016 Collection 1 reprocessing campaign

As a result of the major improvements to the cloud detection algorithm, a reprocessing campaign was performed during 2016. Together with the cloud detection algorithm improvements, several other modifications to the data and metadata were included. The reprocessing, applied to all data from 16 October 2013 onwards, was finished in January 2017 and the complete Collection 1 archive has been available since February 2017. Data files are identified as .V101 in the filenames.

All data are stored on disk in HDF5 and GeoTiff format for fast access. Furthermore, the intermediate Level 2A data products for the entire Collection at all resolutions were made available to the users.

An extensive comparison between PROBA-V Collection 1 and Collection 0 has recently been performed, the results are reported in Toté et al. (2016). This report is available from the PROBA-V Quality Webpage:
http://proba-v.vgt.vito.be/sites/default/files/Quality/PROBA-V%20Collection%201%20Evaluation.pdf.

Further, the PROBA-V and SPOT-VGT data were inter-compared before and after their reprocessing and the impact on the consistency between the two missions was assessed. A short technical note on the main results is available at
http://proba-v.vgt.vito.be/sites/proba-v.vgt.vito.be/files/comparison_between_spot-vgt_and_proba-v_v1.1_website.pdfhttp://proba-v.vgt.vito.be/sites/default/files/Quality/Comparison_SPOT-VGT_PROBA-V_V1.0.pdf.

More information on the SPOT-VGT Collection 3 data archive reprocessing can be found in Toté et al. (2017), which is available from
http://proba-v.vgt.vito.be/sites/proba-v.vgt.vito.be/files/tote_et_al_spot_vgt_collection_3.pdf.

The changes that were implemented during the PROBA-V reprocessing campaign are described below.

2.5.1. Improved cloud detection algorithm

As already explained in Section 2.2.3.1, a new cloud detection algorithm was developed to improve on major detection issues in the operational algorithm. This change gives the most significant differences between the Collection 1 and Collection 0 data products.

2.5.2. Updates to the radiometric ICP files

With changes to the radiometric ICP files, users benefit from improved reflectance values due to updated absolute calibration coefficients, a better inter-camera consistency, and an overall improvement of the radiometric pixel quality.

The changes that were made to the radiometric ICP files include:

1. Inter-camera adjustments to the VNIR absolute calibration coefficients;
2. The application of a degradation model to the SWIR absolute calibration coefficients;
3. Improvement of the low frequency multi-angular coefficients (i.e., equalization) for the SWIR strips of the CENTER camera;
4. Changes to the dark current values;
5. Minor changes to the status of bad pixels.

Users that are interested in more details on the above mentioned changes are referred to Appendix C.

2.5.3. Update product metadata for Climate Forecast (CF) compliancy

Metadata from 16 October 2013 to present were made compliant to the Climate and Forecast metadata conventions (CF v1.6). In Collection 0, metadata were already compliant to these conventions for data from 6 January 2016 onwards. More details on the CF conventions can be found in Section 4.5.

2.5.4. Processing algorithm bug fixes

The following bugs in the processing algorithm were fixed:

1. Bug fix to limit the impact of on-board compression errors for all data that were impacted in Collection 0 before 16 July 2015.
2. Bug fix in the Processing Facility component to the module that checks the satellite attitude data. Some Collection 0 data before 10 February 2016 might be erroneously flagged as 'No data'.

2.6. Antarctica data

As a result of large user interest, it was decided to extend the PROBA-V observations over the Antarctica continent and the surrounding sea ice during the southern hemisphere summer season (November through February) and to provide these observations to the users as Level2A (Top-of-Atmosphere reflectance) products. The data are provided in Polar Stereographic Projection at 100 m, 300 m, and 1 km resolution. An example is shown in Figure 16.

It is noted that the cloud detection algorithm is less accurate over the Antarctic continent, as a result of both the low solar elevation at these high latitudes and the high-reflective snow- and ice-covered surface. Further, similar to the northern hemisphere observations, no atmospheric correction is applied, as the SMAC algorithm becomes increasingly inaccurate for high solar and viewing zenith angles.

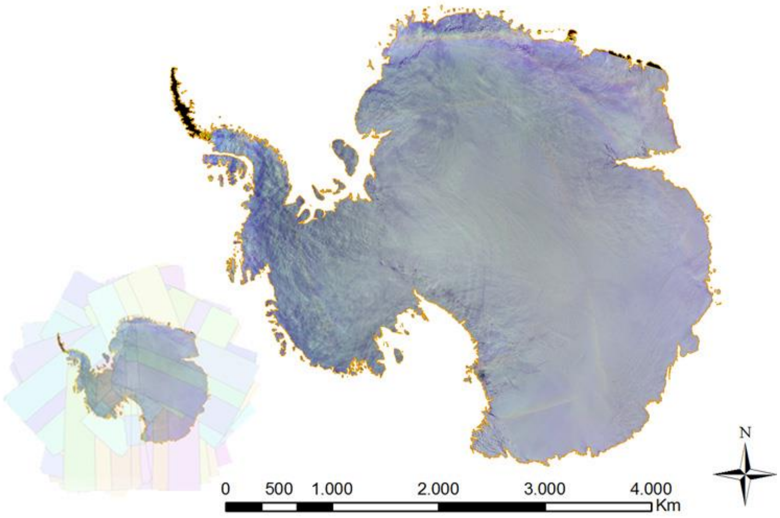


Figure 16: Example of the PROBA-V Level 2A Antarctica data product. The small image on the left indicates the number of orbits that were required to cover the entire continent. Note that the data products also contain observations over the surrounding sea ice, which are not visualised in this example.

3. Product data access

Proba-V Collection 1 data are no longer available to users. For more information on Proba-V Collection 2 product data access, see https://proba-v.vgt.vito.be/en.

4. Data and metadata formats

4.1. HDF5 EOS File Format

PROBA-V data products are disseminated as HDF5 files (Hierarchical Data Format, Version 5, for more information see http://www.hdfgroup.org/HDF5/), which comprises a set of file formats and libraries designed to store and organize large amounts of numerical data. The structure within an HDF5 file has mainly two major object types:

- Datasets, which are multi-dimensional arrays of a homogeneous type
- Groups, which are container structures that can contain other datasets and groups

The HDF5 file format is hierarchical and is built up like a file system. See for example Figure 17, which shows the various Datasets and Groups for a PROBA-V Level 1C file, as well as the BLUE and RED bands opened as images. In HDF5, attributes with additional information are attached to the Datasets and Groups.

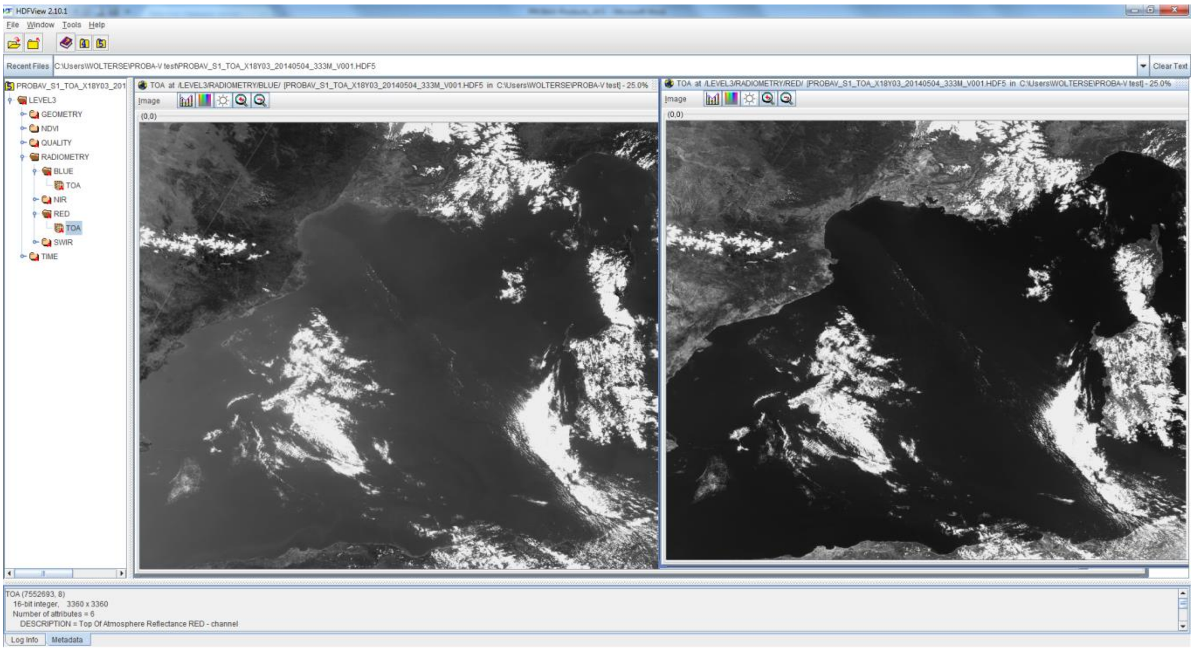


Figure 17: Overview of the Datasets, Groups, and images of the BLUE and RED spectral bands within a PROBA-V Level 1C HDF5 file.

4.1.1. SZIP compression

The HDF5 files are produced using the SZIP (de)compression software. SZIP is a stand-alone software library that ensures lossless compression of scientific data and is superior in both compression rate and (de)compression times during I/O as compared to e.g. GZIP.

Most software packages that can read HDF5 files have the SZIP library included. However, users are referred to the following links to obtain more detailed information on the SZIP performance and to download the SZIP library pre-compiled source code when necessary:

http://www.hdfgroup.org/doc_resource/SZIP/
https://www.hdfgroup.org/downloads/hdf5/

4.1.2. Customization tool

Upon various user requests for extended PROBA-V data tools, VITO developed the PROBA-V Product Customization Tool (PPT). The PPT software facilitates the use of PROBA-V products by providing following customization operations:

- File format conversion (HDF5 → HDF4 / GeoTiff)
- Map projections
    - Albers Equal Area
    - Lambert Equal Area
    - Equi-rectangular
    - Geographic
    - Hammer
    - Sinusoidal
    - Goode Homolosine
    - Interrupted Goode Homolosine
    - Lambert Azimuthal
    - Lambert Conformal Conic
    - Mercator
    - Mollweide
    - Polar Stereographic
    - Stereographic
    - Transverse Mercator
    - Universal Transverse Mercator
- Mosaicking
- Band extraction
- Clipping

The PPT software is compiled for use on multiple operating systems (Unix and Windows) for both 32- and 64-bit systems. An extensive PPT User Manual is included in the download package, available from the PPT download page (http://www.vito-eodata.be/PDF/image/news/PROBA-V_Product_Customization_Tool.html).

4.2. GeoTiff format

The GeoTagged Image File Format (GeoTiff) is a metadata standard that allows for including georeferencing information (ellipsoids, projection, datums, etc.) to a TIFF raster file. The GeoTiff format has become the standard format for most GIS applications, including Quantum GIS, ArcGIS, ERDAS Imagine, etc. GeoTiff images can be properly read by any program/script that is built on the Geospatial Data Abstraction Library (GDAL).

After ordering, the user receives per product date the following GeoTiff files, see Table 5. Please note that this Table shows the filenames for synthesis files. The TIME grid and NDVI data are not available for Level 2A data, as these data are Top-of-Atmosphere segments.

Table 5: PROBA-V GeoTiff filenames and content.

Filename	Content
*RADIOMETRY.tif	Spectral reflectances: BLUE, RED, NIR, and SWIR Band 1: RED Band 2: NIR Band 3: BLUE Band 4: SWIR
*GEOMETRY.tif	Geometry data: SAA, SZA, VNIR VAA, VNIR VZA, SWIR VAA, and SWIR VZA Band 1: SZA Band 2: SAA Band 3: SWIR VAA Band 4: SWIR VZA Band 5: VNIR VAA Band 6: VNIR VZA
*SM.tif	Status Map information
*TIME.tif	Time Grid data
*NDVI.tif	Normalized Difference Vegetation Index

4.3. Algorithm Version Information

The PROBA-V product version number in the filename has three digits, which consists of two parts: the first digit indicates the collection number, with ‘0’ for Collection 0 and ‘1’ for Collection 1. The second and third digit together represent a processing counter for the number of iterations a certain collection has taken till completion. Generally, these numbers will not change once a Collection has been completed and for Collection 1 the version numbering is ‘V101’ for the large majority of product files. However, for Collection 1 also a version number ‘V102’ exists for data between 1 September 2016 and 6 December 2016. This additional reprocessing was necessary due to an incorrect handling of the geometric ICP files in this period.

The various algorithms within the processing chain have an irregular change in versioning. Table 6 presents these algorithms, while Table 7 contains information on the algorithms’ version numbers as per August 2016. The latter information is found in the Level 3 Group metadata attributes.

Table 6: Definition of the various PROBA-V processing algorithms.

Metadata Field	Description
PROCESSINGINFO_GEOMODELLING	Identifier for the algorithm and version of the geometric processing step.
PROCESSINGINFO_RADIOMODELLING	Identifier for the algorithm and version of the radiometric processing step.
PROCESSINGINFO_MAPPING	Identifier for the algorithm and version of the projection step.
PROCESSINGINFO_MOSAIC	Identifier for the algorithm and version of the mosaic processing step.
PROCESSINGINFO_CLOUDICESNOWDETECTION	Identifier for the algorithm and version of the cloud/snow/ice detection processing step.
PROCESSINGINFO_ SHADOWDETECTION	Identifier for the algorithm and version of the shadow detection processing step.
PROCESSINGINFO_ATMOSPHERIC_ CORRECTION	Identifier for the algorithm and version of the atmospheric correction processing step.
PROCESSINGINFO_COMPOSITING	Identifier for the algorithm and version of the compositing processing step.

Table 7: Processing algorithm versions for Collection 1 data.

Metadata Field	Value	Description
PROCESSINGINFO_GEOMODELLING	PROBAV_GEOMODELLING_V1.0	Initial version of the geometric modelling algorithm
PROCESSINGINFO_ RADIOMODELLING	PROBAV_RADIOMODELLING_V1.0	Initial version of the radiometric modelling algorithm
PROCESSINGINFO_MAPPING	PROBAV_MAPPING_V1.0	Initial version of the Geometric modelling algorithm
PROCESSINGINFO_MOSAIC	PROBAV_MOSAIC_V1.0	Initial version of the mosaicking algorithm
PROCESSINGINFO_ CLOUDICESNOWDETECTION	PROBAV_ CLOUDICESNOWDETECTION_V2.0	Initial version of the cloud and snow/ice detection algorithm
PROCESSINGINFO_ SHADOWDETECTION	PROBAV_ SHADOWDETECTION_V1.0	Initial version of the shadow detection algorithm
PROCESSINGINFO_ATMOSPHERIC_ CORRECTION	PROBAV_ATMCORR_ SMAC_V1.0	Initial version of the atmospheric correction algorithm
PROCESSINGINFO_COMPOSITING	PROBAV_COMPOSITING_MVC_V1.0 PROBAV_COMPOSITING_MVC_V2.0 PROBAV_COMPOSITING_MVC_V2.1	Initial version of the MVC compositing algorithm Same as PROBAV_COMPOSITING_MVC_V1.0, but with the following changes for the 1 km compositing: • No pixel selection based on viewing zenith angle (VZA) and Solar Zenith Angle (SZA) thresholds • Disabled checking SWIR quality validity This version was applicable for all 1 km synthesis products processed from 13/06/2014 – 15/07/2014. Same as PROBAV_COMPOSITING_MVC_V2.0, but with the following changes: Fixed time grid (minutes since start of the synthesis).

4.4. PROBA-V Product Files Description

4.4.1. Level 1C Product File Naming and Content

The file format for the Level 1C products is as follows:

PROBAV_L1C_*<DATE>*_*<TIME>*_*<CAMERA>*_**V***<VERSION>***.hdf5**

In which:

<DATE>	start date of the segment identifier (format: YYYYMMDD).
<TIME>	Segment start time (hhmmss).
<CAMERA>	camera identifier 1: left camera (descending orbit, east) 2: central camera 3: right camera (descending orbit, west)
<VERSION>	version identifier, three digits starting from '001' for Collection 0 and from '101' for Collection 1

Example: the filename *PROBAV_L1C_20140517_121832_1_V001.hdf5* represents the data that was observed from 17 May 2014 12:18:32 UTC onwards with the LEFT camera.

The segment files contain the following dataset structure:

- **LEVEL-1A**
  This group contains the raw uncompressed digital number per pixel. It also contains the **Platform** information provided by the spacecraft.

- **LEVEL-1B**
  This group gives the geometric processing output. It contains the geographical coordinates (latitude, longitude) for each pixel at heights of 0 and 5000 m above sea level. It also contains the viewing and illumination geometry per pixel.

- **LEVEL-1C**
  Contains the radiometrically corrected Top-Of-Atmosphere reflectance value per pixel. It also contains a quality indicator, which gives information per pixel on the reliability of the value.

The above Levels have the following datasets:
- BLUE
- NIR
- RED
- SWIR1
- SWIR2
- SWIR3

The respective datasets contain the Digital Number counts for the respective spectral bands, with attributes providing information on the scale and offset values required to convert the DN to physical reflectance values.

Figure 18 displays the dataset structure of an Level 1C file. Further, each of the Levels contains additional HDF5 Attribute Tables in which detailed information on geolocation, processing, etc. is stored. Detailed explanations of the entire dataset structure for the Level 1C files are given in Appendix E1.

Figure 18: Dataset structure of a Level 1C product file.

4.4.2. Level 2A Product File Naming and Content

The file format for Level 2A product files is as follows:

PROBAV_L2A_*<DATE>***_***<TIME>*_*<CAMERA>***_***<RESOLUTION>*_**V***<VERSION>***.hdf5**

In which:

<DATE>	start date of the segment identifier (format: YYYYMMDD).
<TIME>	start time (UTC) of the segment (format: hhmmss).
<CAMERA>	camera identifier 1: left camera (descending orbit, east) 2: central camera 3: right camera (descending orbit, west)
<VERSION>	version identifier, three digits starting from ‘001’ for Collection 0 and from ‘101’ for Collection 1.

Example: the filename *PROBAV_L2A_20160210_105508_1_1KM_V001.HDF5* represents data that were collected from observations using the left camera, starting 10 February 2016 at 10:55:08 UTC.

Figure 19 shows the dataset structure of a Level 2A HDF5 file.

Figure 19: Dataset structure of a Level 2A product file.

The dataset structure is built around the Level 2A Main Group. Within this Main Group, the following Groups can be distinguished:

- **GEOMETRY**
  Contains the viewing and illumination geometry per pixel.

- **QUALITY**
  Contains a quality indicator per pixel, consisting of an observation indicator (clear, cloud, ice, shadow, undefined), a land/sea flag, and a radiometric quality indicator.

- **RADIOMETRY**
  Contains the reflectance value per pixel at Top-of-Atmosphere (TOA).

Detailed explanations of the dataset structure, dataset contents, and data types for the Level 2A files are given in Appendix E2.

4.4.3. Synthesis Product File Naming and Content

The file format for the synthesis products (S1, S5, and S10) is as follows:

PROBAV_<*TYPE*> **_**<*TILE-ID*>**_**<*DATE*>**_**<*GRID*>**_V**<*VERSION*>**.hdf5**

In which:

<TYPE>	product type (‘S1_TOA’, ‘S1_TOC’ or ‘S10_TOC’)
<TILE ID>	tile identifier. ‘X00Y00’ is the identifier for the top-left tile, numbering increases in eastward and southward direction for X and Y, respectively. See Figure 20 for the tile numbering.
<DATE>	start date of the synthesis (YYYYMMDD)
<GRID>	spatial resolution (300 m or 1 km)
<VERSION>	version identifier, three digits starting from ‘001’ for Collection 0 and from ‘101’ for Collection 1.

Example: The top-left tile of the third 1 km S10 of September 2013 has filename: **PROBAV_S10_TOC_X00Y00_20130921_1KM_V001.hdf5.**

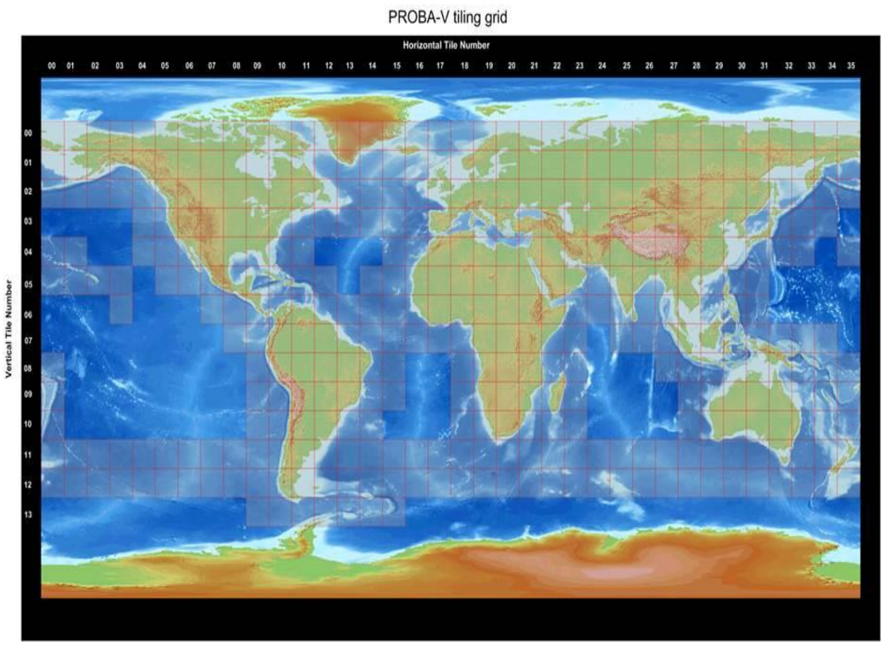


Figure 20: PROBA-V tile numbering.

Figure 20 explains the tile numbering (tiles have 10° × 10° dimensions). The top-left tile is numbered 'X00Y00' (having top-left coordinates 180° E, 75° N), with the X and Y tile numbers increasing eastward and southward, respectively.

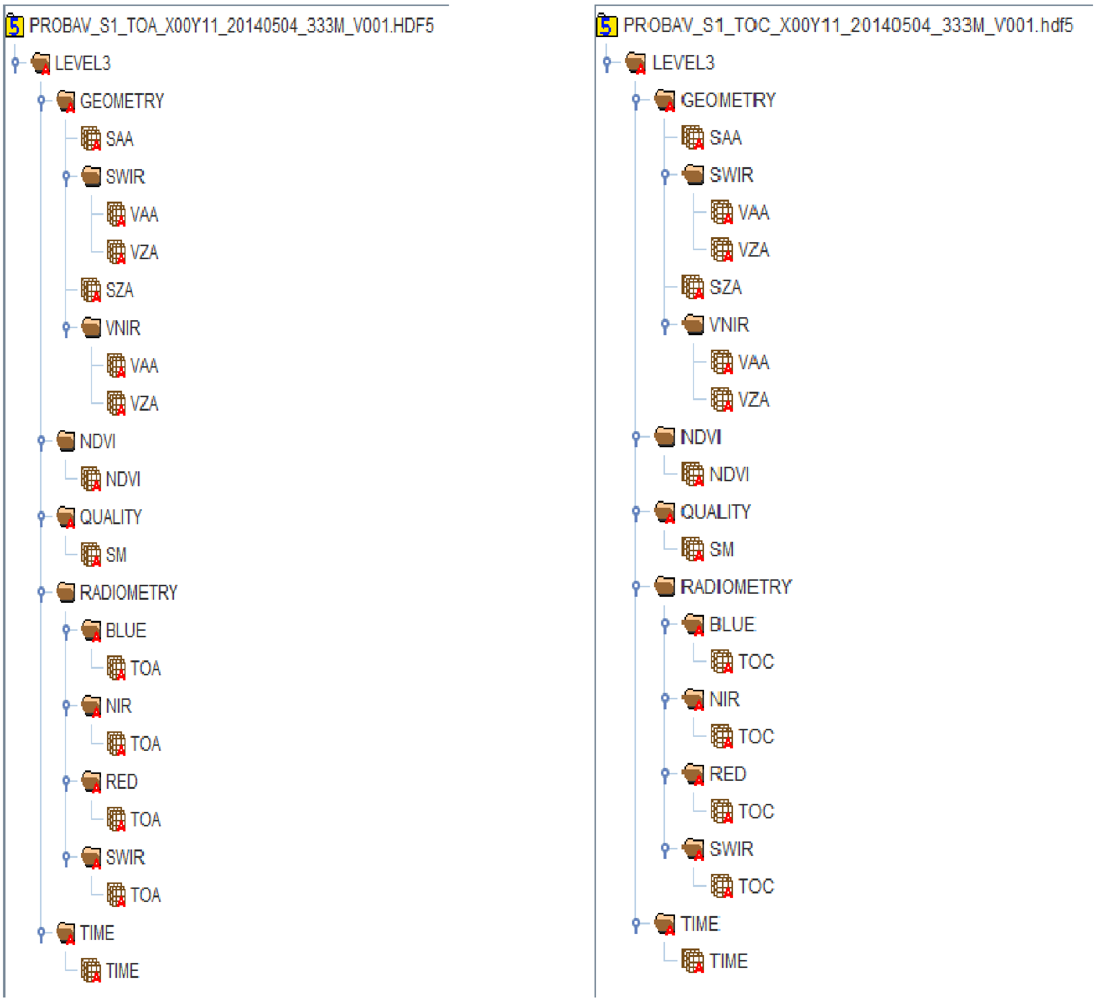


Figure 21: Dataset structure of S1 TOA (left) and TOC (right) product files.

The dataset structure is built around the Level 3 Main Group. Within this Main Group, the following Groups can be distinguished:

- **GEOMETRY**
  Contains the viewing and illumination geometry for each product pixel.

- **NDVI**
  Contains the Normalized Difference Vegetation Index (NDVI) for each product pixel.
  The NDVI is calculated from the RED and NIR TOA (S1 TOA) or TOC (S1 TOC) reflectances:

$$NDVI = \frac{R_{NIR} - R_{RED}}{R_{NIR} + R_{RED}} \quad \text{(Eq. 1)}$$

The conversion from digital values in the synthesis files to physical values is done as follows:

$$PV = (DN - offset)/scale \qquad \text{(Eq. 2)}$$

In which PV is the physical value and *DN* is the digital value. For NDVI, the *offset* and *scale* values are 20 and 250, respectively, giving a valid physical range of [-0.08, 0.92] for digital value range of [0, 250]. Observations for which the NDVI underruns or exceeds the lower or upper limits are given DN values 0 and 250, respectively. An overview of scale, offset, and no data values for all dataset types is provided in Section 4.6.1.

- **QUALITY**
  Contains a quality state indicator per pixel, consisting of an observation indicator (clear, cloud, ice, shadow, undefined), a land/sea flag and a radiometric quality indicator.

- **RADIOMETRY**
  Contains the TOA or TOC reflectance value per pixel.

- **TIME**
  Contains the date and time of observation, expressed as the number of minutes since the beginning of the synthesis period in UTC. It is noted that the value for ‘no data’ in the TIME grid is set to ‘0’, which can also be a real value. Therefore users should check the Status Map for additional information to ascertain whether the TIME grid pixel carries indeed a ‘no data’ value.

Figure 21 presents the entire dataset structure for the TOA and TOC Synthesis products. Detailed information on the Groups, Attributes, and Datasets is given in Appendix E3.

4.5. Climate Forecast (CF) compliant metadata

PROBA-V HDF5 Level 3 data (i.e., the S1, S5, and S10 synthesis files) product metadata have been extended with Climate and Forecast (CF) v1.6 (CF-1.6) compliant metadata. In Collection 0, data from 6 January 2016 onwards contained CF compliant metadata and this was extended to the entire archive (from 16 October 2013 onwards) for Collection 1.

The CF metadata are designed to promote the processing and sharing of files created with the NetCDF Application Programming Interface (API) and are increasingly accepted and adopted as a primary metadata standard. The conventions define metadata that provide a definitive description of what the data in each variable represents, as well as the spatial and temporal data properties and units. This enables users from different sources to decide which quantities are comparable and facilitates building applications with powerful extraction, regridding, and display capabilities. An overview of the CF-1.6 compliant metadata attributes that are found in the PROBA-V Level 3 products is given in Table 8. A detailed overview of the CF metadata convention is available at http://cfconventions.org/.

Table 8: Explanation of the CF v1.6 compliant metadata attributes.

Metadata location	Attributes	Explanation
<root>	Conventions _CoordSysBuilder Institution source	Global CF convention attribute. Used Class of CF metadata Name of the institute that produced the product. The production method of the original data.
./crs	GeoTransform _CoordinateAxisTypes _CoordinateTransformType grid_mapping_name inverse_flattening longitude_of_prime_meridian semi_major_axis spatial_ref	Pixel/line to coordinate space. transformation coefficients Label used in GDAL library. Transformation type. Name used to identify the grid mapping. Inverse flattening of the ellipsoid associated with the geodetic datum. Zero degree longitude. Major axis (in meters) of the ellipsoid. Spatial reference system in the Open Geospatial Consortium (OGC) Well-Known Text (WKT) format.
./lat and ./lon	_CoordinateAxisType axis long_name standard_name units	Label used in GDAL library. Axis indicator Descriptive name indicating the variable's content. CF v1.6 standardized name that references the variable's content. The latitude/longitude unit.
LEVEL3/<GROUP>/<DATASET>	_FillValue add_offset grid_mapping long_name scale_factor standard_name units calendar (only for TIME attribute)	Value representing missing or undefined data, must be outside of valid data range. Reference to offset value used in the conversion from digital to physical values. Reference to the grid mapping value. Descriptive name indicating a variable's content. Reference to scale factor used in the conversion from digital to physical values. CF v1.6 standardized name that references the variable's content. Units of the variable's content, taken from Unidata's 'udunits' library.

		Information on the calendar system (Gregorian, Julian, or combined).

4.6. Data viewing and handling

The HDF5 file format is readable for most data interpretation languages, such as IDL, R, and Python. Further, applications exist to quickly view the data as images and to perform basic calculations on the data. Examples of such applications are HDFView and Quantum GIS (QGIS).

4.6.1. DN to PV value scaling

The reflectances that are provided in the Level 1C, Level 2A, and Level 3 data files are presented as Digital Count Numbers (DN). In order to obtain reflectance values a conversion is necessary, which is done using the formula below:

$$PV = (DN - OFFSET) / SCALE$$

With DN the Digital Count Number and PV the Physical Values, respectively. The *SCALE* and *OFFSET* parameters can be found in the various dataset attributes (reflectances, zenith and azimuth angles, NDVI, Status Map, and Time grid) in the Level 1C, Level 2A, and Level 3 files. Table 9 contains the scale, offset, and no data values for each of these dataset types. Note that for the Level 1C files the SWIR channel data contain the observations for each of the three strips.

Table 9: Scale, offset, and no data values for the PROBA-V dataset types.

Product type	Scale	Offset	No data
Reflectances	2000	0.0	-1.0
Azimuth Angles	0.66667	0.0	255.0
Zenith Angles	2.0	0.0	255.0
NDVI	250.0	20.0	255.0
Status Map	1.0	0.0	2.0
Time	1.0	0.0	0.0

4.6.2. Opening HDF5 S1 and S10 in ENVI 5.2

From ENVI 5.2 onwards, PROBA-V HDF5 files are supported. Opening the HDF5 files is done as follows:

- Click ‘File’ → ‘Open As’ → ‘PROBA-V’ → ‘Synthesis (S1 & S10)’, see Figure 22.
- The VNIR bands (RED, NIR, BLUE) of the PROBA-V HDF5 file are opened.

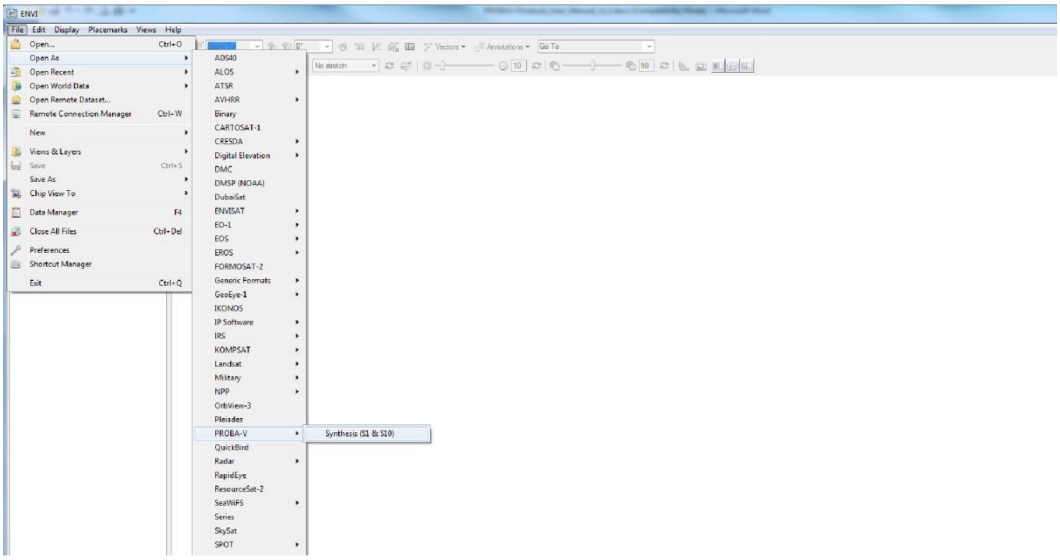


Figure 22: Opening a PROBA-V HDF5 file in ENVI 5.2.

For further details on the image processing using ENVI 5.2, please consult the ENVI documentation (https://www.harrisgeospatial.com/docs/using_envi_Home.html).

4.6.3. Opening HDF5 in Interactive Data Language (IDL)

The example program below shows how to open a PROBA-V S1 synthesis HDF5 file in IDL and how to read the data. In all examples, commented code is highlighted in red.

```
PRO read_hdf5
    ;Open the HDF5 file.
    file = 'PROBAV_S1_TOA_X11Y07_20140607_300 m_V001.hdf5'
    file_id = H5F_OPEN(file)

    ;Open the image dataset within the file.
    dataset_id1 = H5D_OPEN(file_id, '/RADIOMETRY/BLUE/TOA')

    ;Read in the actual image data.
    image = H5D_READ(dataset_id1)

    H5D_CLOSE, dataset_id1
    H5F_CLOSE, file_id
end
```

4.6.4. Opening HDF5 in R

The example program below shows how to open a PROBA-V S10 synthesis HDF5 file and how to read the data in R, using the `h5r` package.

```
read_hdf5 <-function(){
    require(h5r)

    filename <- "PROBAV_S1_TOA_X11Y07_20140607_300 m_V001.hdf5"

    #extract the HDF5 dataset
    h5 <- H5File(filename,'r')
    dblu  <- getH5Dataset(h5, "/LEVEL3/RADIOMETRY/BLUE/TOA")

    #get the image values and store into 3360 x 3360 matrix
    blue  <- array(readH5Data(dblu), c(3360, 3360))
}
```

Note that alternative packages (such as `rhdf5`) exist; the syntax to open and read the HDF5 file will be slightly different. An example script using `rhdf5` is shown below and evidently data are more conveniently extracted using this package.

```
read_hdf5_alt <-function(){
    require(rhdf5)

    filename <- "PROBAV_S1_TOA_X11Y07_20140607_300 m_V001.hdf5"

    #extract the TOA reflectances for the four spectral bands
    d_red <- h5read(h5file, "LEVEL3/RADIOMETRY/RED/TOA") / 2000
    d_nir <- h5read(h5file, "LEVEL3/RADIOMETRY/NIR/TOA") / 2000
    d_blu <- h5read(h5file, "LEVEL3/RADIOMETRY/BLUE/TOA") / 2000
    d_swi <- h5read(h5file, "LEVEL3/RADIOMETRY/SWIR/TOA") / 2000
}
```

4.6.5. Opening HDF5 in Python

The example program below shows how to open a PROBA-V S1 synthesis HDF5 file in Python, using the `h5py` and `numpy` packages.

```
#Import h5py library
Import h5py

#Open HDF5 file with h5py, read-write mode
h5f=h5py.File ('PROBAV_S1_TOA_X11Y07_20140607_300 m_V001.hdf5','r+')

#Use Python dictionary syntax to explore the HDF5 structure
h5f.keys()

#Get dimensions
h5f['/LEVEL3/RADIOMETRY/BLUE/TOA'].shape

#Get data type
h5f['/LEVEL3/RADIOMETRY/BLUE/TOA'].dtype
```

```
#Get value array
toa_b = h5f['/LEVEL3/RADIOMETRY/BLUE/TOA'].value
```

4.6.6. Opening HDF5 in HDFView

The example below shows how to open a PROBA-V S1 synthesis HDF5 file in HDFView (v2.10). After starting up HDFView, select a file, see Figure 23.

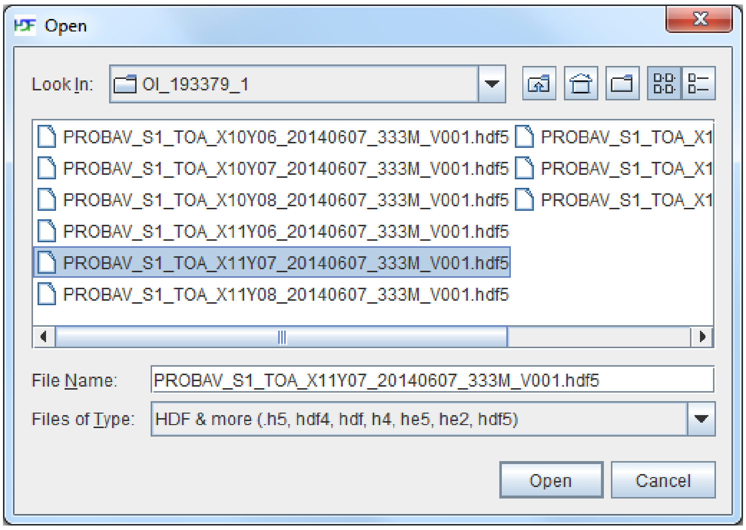


Figure 23: Dialog box for opening an HDF5 file in HDFView.

Upon opening the HDF5 file, the selected band data (in this case the BLUE band) can be viewed through right-clicking the band. Select 'Open As' (see Figure 24), which gives 2 options: open the data as a spreadsheet or view the data as an image.

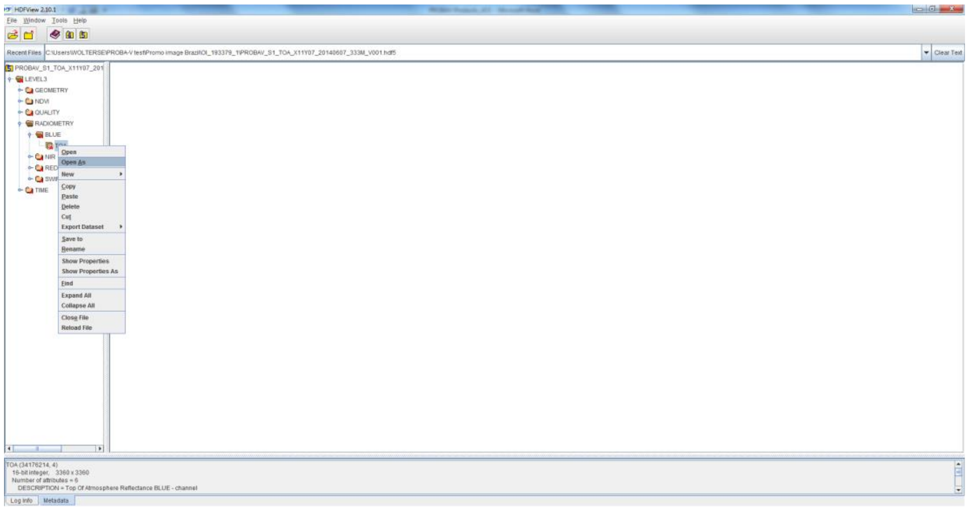


Figure 24: Band selection to open the dataset as either spreadsheet or image.

When selecting 'Open As Image', the user has to choose a color palette, see Figure 25. After clicking 'OK', the dataset is loaded and presented as an image. For further options and basic statistics to be calculated within HDFView, reference is made to the HDFView User Guide:

http://www.hdfgroup.org/products/java/hdfview/UsersGuide/

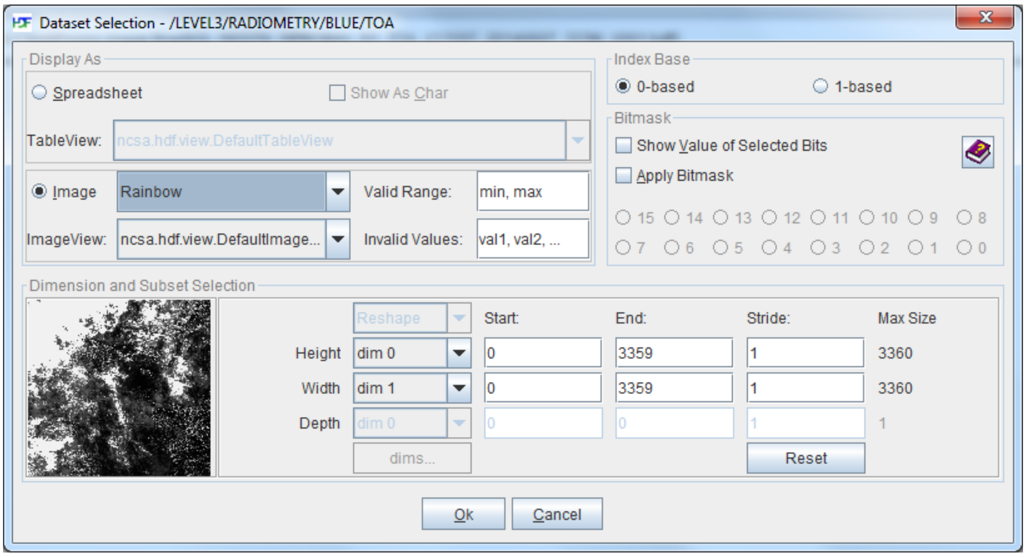


Figure 25: Colour palette selection for image viewing.

4.6.7. Opening HDF5 in Quantum GIS

Below a short description on how to open a PROBA-V HDF5 file in Quantum GIS (QGIS, version 2.16 Nødebo) is given.

In the Quantum GIS main screen, select the 'Open Raster' icon. A dialog box to select the raster file is opened, see Figure 26. Once the HDF5 file is opened, another dialog box for selecting one or more bands is opened, see Figure 27. In this example the BLUE band is selected for further viewing. After band selection, the proper coordinate reference system needs to be chosen, which is presented in Figure 28. Once these three steps have been completed, the band image is loaded and further actions can be performed.

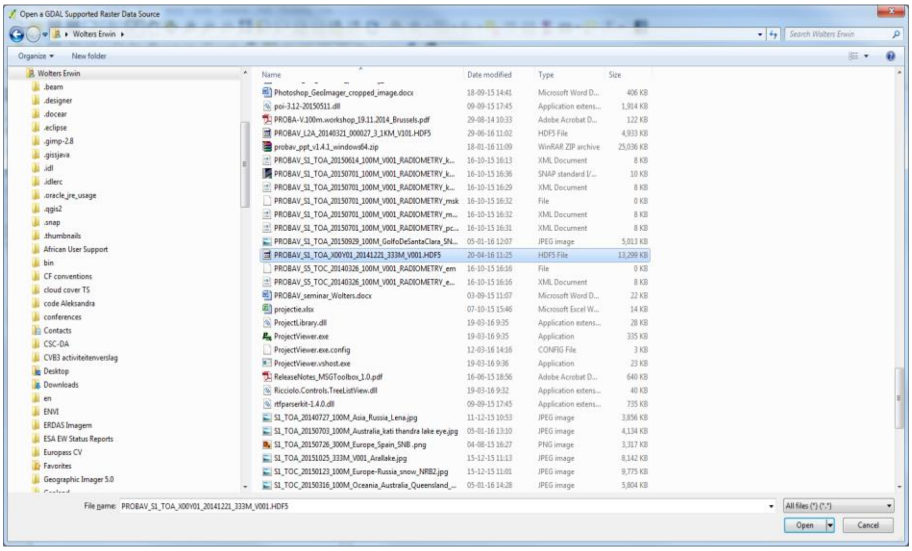


Figure 26: Dialog box for opening a raster file in QGIS 2.16.

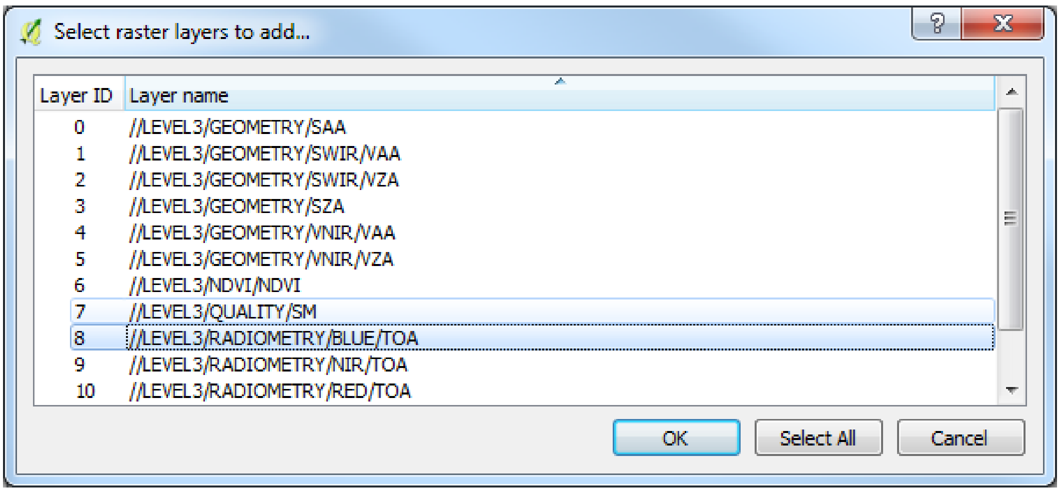


Figure 27: Selection of the BLUE band TOA data.

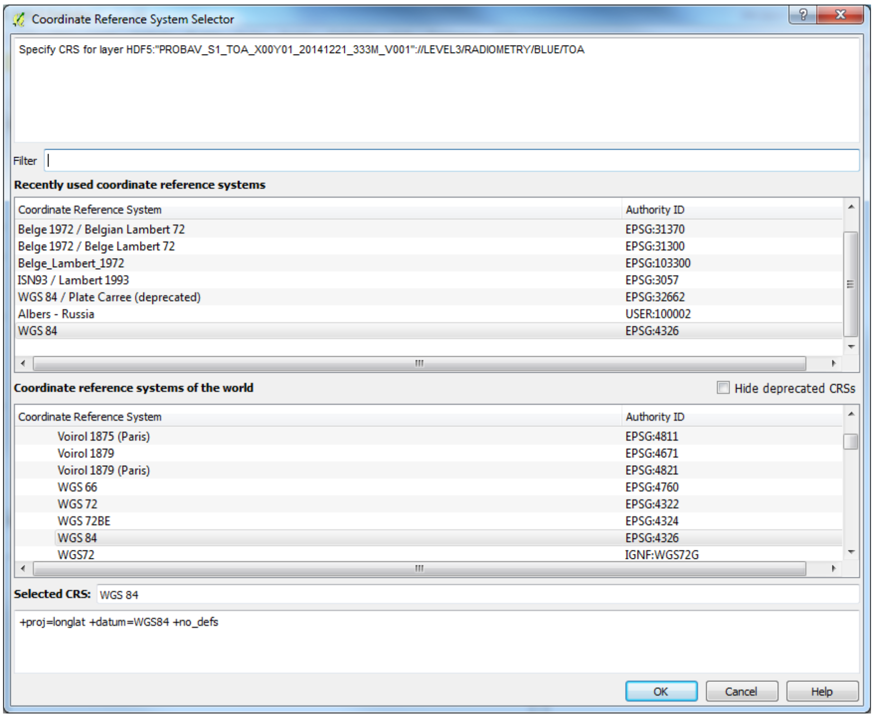


Figure 28: Selection of the WGS84 Coordinate Reference System (CRS).

It is noted that, although a Coordinate Reference System (CRS) is selected, a discrepancy between the QGIS and HDF5 Geotie point definitions exists. This will result in the HDF5 image being displayed with image coordinates, rather than with the WGS84 lat/lon coordinates. In order to obtain proper overlays with e.g. vector files, users are advised to download PROBA-V images in GeoTiff format or to open the HDF5 in GIS programs that are compliant with the HDF5 Geotie point definition (e.g. ENVI 5.2).

4.6.8. Opening HDF5 in SPIRITS

The 'Software for the Processing and Interpretation of Remotely sensed Image Time Series' (SPIRITS) package (Eerens and Haesen, 2014) was developed at VITO and is a Windows-based software tool aiming at the analysis of remotely sensed earth observation data. Although it includes a wide range of general purpose functionalities, the focus is put on the processing of image time series, derived from various low- and medium-resolution sensors.

SPIRITS enables conversion of the PROBA-V HDF5 file format into the ENVI file format upon further viewing and processing. More details on the conversion and the SPIRITS software can be found in the SPIRITS Manual: http://spirits.jrc.ec.europa.eu/files/SpiritsManual_150.pdf

4.6.9. MATLAB PROBA-V reader

Given the need expressed by a number of users to be able to conveniently read, extract, and convert PROBA-V HDF5 data using MATLAB, a MATLAB reader was developed at VITO. The MATLAB reader enables users to read and convert PROBA-V S1 and S10 data files into e.g. standard ENVI format. The reader can be downloaded from http://proba-v.vgt.vito.be/en/s1s10-matlab-reader-tool, additional information and a video tutorial is also available.

4.6.10. Sentinel Application Platform (SNAP) PROBA-V Toolbox

The Sentinel Application Platform (SNAP) comprises a set of Toolboxes that facilitate the analysis, processing, and visualisation of the Sentinel-1, -2, and -3 satellite data. As part of the Sentinel Toolboxes, a first version of the PROBA-V Toolbox was released in January 2017. The PROBA-V Toolbox enables users to conveniently access, analyse, and visualise PROBA-V Level 2A and Level 3 data.

The Sentinel Toolboxes can be downloaded from http://step.esa.int/main/download/ and the PROBA-V Toolbox v1.0 release notes are available at https://github.com/senbox-org/probavbox/blob/1.x/ReleaseNotes.md.

5. Quality assurance

Both the Segment and Synthesis product files are delivered with quality indicators. Below these indicators are shortly explained. Reference is made to Appendices E1 – E4 for detailed descriptions of the Segment and Synthesis metadata.

5.1. Level 1C files

For Level 1C files, the quality is indicated by the Q Dataset, which is located in the LEVEL 1C STRIP Group (see Appendix E1 for more details). The pixel quality for the Level 1C data is decoded as 8-bit unsigned integers, the values and their meaning are given in Table 10.

Table 10: Explanation of the pixel quality indicators in the Segment Product.

Status	Explanation
0	'Correct': no quality issues encountered
1	'Missing': the pixel value is missing due to a bad detector
2	'WasSaturated': the DN value of the pixel is equal to 4095 (coded in 12 bits =2^12-1)
3	'BecameSaturated': during the calculation of the TOA reflectance the value became higher than 4095
4	'BecameNegative': during the calculation of the TOA reflectance the value became lower than 0
5	'Interpolated': the value of the pixel radiance was interpolated using the neighbouring pixels
6	'BorderCompressed': the pixel quality is uncertain due to on-board compression artefacts

5.2. Level 2A and synthesis product files

In the Level 2A, S1, S5, and S10 product files, the quality indicator is located in the SM (Status Map) Dataset within the QUALITY Group. The SM Dataset contains a quality state indicator per pixel, consisting of an observation indicator (clear, cloud, ice, shadow, undefined), a land/sea flag, and a radiometric quality indicator. Table 11 lists the various quality values. Please note that bits 8 – 11, containing additional information on the observational coverage for each band, are only available for Level 2A data.

Table 11: Explanation of the pixel quality indicators in the Status Map Dataset. Bits indicated with an asterisk are only available for Level2A data.

Bit (LSB to MSB)	Description	Value	Key
0-2	Cloud/Ice Snow/Shadow Flag	000 001 010 011 100	Clear Shadow Undefined Cloud Ice
3	Land/Sea	0 1	Sea Land
4	Radiometric quality SWIR flag	0 1	Bad Good
5	Radiometric quality NIR flag	0 1	Bad Good
6	Radiometric quality RED flag	0 1	Bad Good
7	Radiometric quality BLUE flag	0 1	Bad Good
8*	SWIR coverage	0 1	No Yes
9*	NIR coverage	0 1	No Yes
10*	RED coverage	0 1	No Yes
11*	BLUE coverage	0 1	No Yes

The Status Map information can be easily read in most programming languages. Below we present how to read and extract the various flags for bits 0 – 2 (see Table 11) from the Status Map Level 2A data for R and Python.

R

The ‘`rhdf5`’ package is used to extract the Status Map dataset from the Level 2A file, while ‘`R.utils`’ contains the ‘`intToBin`’ function, which is used to convert the Status Map integer values to binary. Comments are highlighted in red.

```
require(rhdf5)
require(R.utils)

L2_file     <- 'C:\PROBAV_L2A_20150506_085613_3_1KM_V001.HDF5'
SM          <- h5read(L2_file, "LEVEL2A/QUALITY/SM")

#convert the Status Map integers to binary values
b_SM <- intToBin(SM)

#extract the 3 rightmost bits from 12 (in Level 2A, 8 bits in Level 3 data)
and assign clear, shadow, undefined, cloudy, and snow/ice
clr  <- which(substr(b_SM, 10, 12) == "000")
shw  <- which(substr(b_SM, 10, 12) == "001")
und  <- which(substr(b_SM, 10, 12) == "010")
```

```
cld  <- which(substr(b_SM, 10, 12) == "011")
ice  <- which(substr(b_SM, 10, 12) == "100")
```

Python

In Python, the `h5py` library conveniently handles the opening, reading, and extraction of HDF5 datasets. The `numpy` library in this example contains the '`where`' statement, which returns those array indices that fulfill a certain criterion. Bitwise operations are used here, with '`&1`', '`&2`', and '`&4`' indicating evaluations on the 3 least significant bits.

```
Import h5py
Import numpy as np

L2_file = 'C:\PROBAV_L2A_20150506_085613_3_1KM_V001.HDF5'
h5 = h5py.File (L2_file,'r+')
SM = h5['/LEVEL2A/QUALITY/SM'].value

#Evaluate the three least significant bits for 'clear', 'shadow', 'undefined', 'cloud', and 'snow/ice' and assign the outcome to variables
clr = np.where((SM&1 == 0) & (SM&2 == 0) & (SM&4 == 0))
shw = np.where((SM&1 != 0) & (SM&2 == 0) & (SM&4 == 0))
und = np.where((SM&1 == 0) & (SM&2 != 0) & (SM&4 == 0))
cld = np.where((SM&1 != 0) & (SM&2 != 0) & (SM&4 == 0))
ice = np.where((SM&1 == 0) & (SM&2 == 0) & (SM&4 != 0))
```

5.3. PROBA-V Quality Webpage

A Quality Webpage is available at http://proba-v.vgt.vito.be/en/quality/quality and provides information on the Quality Assessment and the various methods applied to maintain PROBA-V's data quality at the highest possible level, from the raw satellite observations through the value-added products available at the PROBA-V Data Portal. The following topics are highlighted:

- Calibration
  - Introduction
  - Radiometric calibration
  - Geometric calibration
- Platform status
  - Geolocation accuracy
  - Spectral Response Functions
  - Quarterly Image Quality Reports
- Product evaluation
  - Introduction
  - Access to PROBA-V and SPOT-VGT Evaluation Reports
- Product and Algorithm Information
  - PROBA-V reference documentation
- Known issues
- PROBA-V Quality Working Group
  - Introduction, objectives, and composition

- Members
- Contact Point

REFERENCES

Ackerman, S., R. Frey, K.Strabala, Y. Liu, L. Gumley, B. Baum, and P. Menzel (2010), Discriminating clear-sky from cloud with MODIS — Algorithm Theoretical Basis Document. *Products: MOD35. ATBD Reference Number: ATBD-MOD, 6, Cooperative Institute for Meteorological Satellite Studies (CIMSS), University of Wisconsin – Madison,* 117 pp.

Dierckx, W., S. Sterckx, I. Benhadj, S. Livens, G. Duhoux, T. Van Achteren, M. Francois, K. Mellab, G. and Saint, (2014), PROBA-V mission for global vegetation monitoring: standard products and image quality. *Int. J. Remote Sens*, **35**, 2589 – 2614, DOI: 10.1080/01431161.2014.883097.

Eerens, H. and D. Haesen, (2014). SPIRITS Manual, 322 pp, available at http://spirits.jrc.ec.europa.eu/, 369 pp.

Francois, M., S. Santandrea, K. Mellab, D. Vrancken, and J. Versluys (2014), The PROBA-V mission: The space segment. *Int. J. Remote Sensing*, **35**, 2548 – 2564, doi: 10.1080/01431161.2014.883098.

Goor, E., J. Dries, D. Daems, M. Paepen, F. Niro, P. Goryl, P. Mougnaud, and A. Della Vecchia, (2016), PROBA-V Mission Exploitation Platform. *Remote Sens.* **8**, 564, doi:10.3390/rs8070564.

Holben, B. N. (1986), Characteristics of maximum-value composite images from temporal AVHRR data. *Int. J. Remote Sens.*, **7**, 1417 – 1434.

Jedlovec, G. (2009), Automated detection of clouds in satellite imagery. Book chapter (ISBN: 978-953-307-005-6): Advances in geoscience and remote sensing. Available from http://weather.msfc.nasa.gov/sport/journal/pdfs/2009_GRS_Jedlovec.pdf.

Lissens, G., P. Kempeneers, F. Fierens, and J. Van Rensbergen (2000), Development of cloud, snow, and shadow masking algorithms for VEGETATION imagery. In *Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International* (Vol. 2, pp. 834 - 836).

Maisongrande, P., B. Duchemin, and G. Dedieu, (2004), VEGETATION/SPOT: an operational mission for the Earth monitoring; presentation of new standard products. *Int. J. Remote Sens.*, **25**, 9 – 14.

Muller, J. P., G. López, G. Watson, N. Shane, T. Kennedy, P. Yuen, and S. Pinnock (2012), The ESA GlobAlbedo Project for mapping the Earth's land surface albedo for 15 Years from European Sensors. In: *Geophysical Research Abstracts* (Vol. 13, p. 10969).

Plummer S., J. Chen, G. Dedieu, and M. Simon, (2003), GLOBCARBON Detailed Processing Model GLBC-ESL-DPM-V1.3, 202 pp.

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

Riazanoff, S., 2004: SPOT1-2-3-4-5 Satellite Geometry Handbook, *GAEL-P135-DOC-001,* **1**, Revision 4, 82 pp.

Simpson, J. J., Z. Jin, and J.R. Stitt (2000), Cloud shadow detection under arbitrary viewing and illumination conditions. *IEEE T. Geosci. Remote*, **38**, 972 – 976.

Stelzer, K., M. Paperin, G. Kirches, and C. Brockmann (2016), PROBA-V cloud mask validation. *Brockmann Consult Validation Report,* 23 pp.

Sterckx, S., I. Benhadj, G. Duhoux, S. Livens, W. Dierckx, E. Goor, S. Adriaensen, W. Heyns, K. Van Hoof, K., G. Strackx, K. Nackaerts, I. Reusen, T. Van Achteren, J. Dries, T. Van Roey, K. Mellab, R. Duca, and J. Zender, (2014), The PROBA-V mission: image processing and calibration. *Int. J. Remote Sens.*, *35*(7), 2565 – 2588, doi: 10.1080/01431161.2014.883094.

Tarpley, J. D., S.R. Schneider, and R.L. Money, (1984), Global vegetation indices from the NOAA-7 meteorological satellite. *J. Clim. Appl. Meteorol*, **23**, 491 – 494.

Toté, C., E. Swinnen, S. Sterckx, D. Clarijs, C. Quang, and R. Maes, (2017), Evaluation of the SPOT/VEGETATION Collection 3 reprocessed dataset: Surface reflectances and NDVI. *Remote Sens. Environ.,* **201**, 219-233.

Toté, C., E. Swinnen, W. Dierckx, S. Sterckx, and D. Clarijs, (2016), Evaluation of the re-processed PROBA-V archive (Phase 2), Interim Report, 74 pp, http://proba-v.vgt.vito.be/sites/default/files/Quality/PROBA-V%20Collection%201%20Evaluation.pdf.

Thuillier, G., M. Hersé, T. Foujols, W. Peetermans, D. Gillotay, P.C. Simon, H. and Mandel (2003), The solar spectral irradiance from 200 to 2400 nm as measured by the SOLSPEC spectrometer from the ATLAS and EURECA missions. *Sol. Phys.*, **214**, 1 – 22, doi: 10.1023/A:1024048429145.

Vancutsem, C. J.-F. Pekel , P. Bogaert and P. Defourny (2007), Mean Compositing, an alternative strategy for producing temporal syntheses. Concepts and performance assessment for SPOT VEGETATION time series, International Journal of Remote Sensing, 28:22, 5123-5141, DOI: 10.1080/01431160701253212, http://dx.doi.org/10.1080/01431160701253212.

Vermote, E. F., D. Tanré, J.L. Deuze, M. Herman, and J.J. Morcette, (1997), Second simulation of the satellite signal in the solar spectrum, 6S: An overview, *IEEE T. Geosci. Remote*, **35**, 675 – 686.

Zhu, Z., and C.E. Woodcock, (2012), Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.*, **118**, 83 – 94, doi:10.1016/j.rse.2011.10.028.

APPENDICES

Appendix A: PROBA-V Spectral Response Functions (SRF) per camera

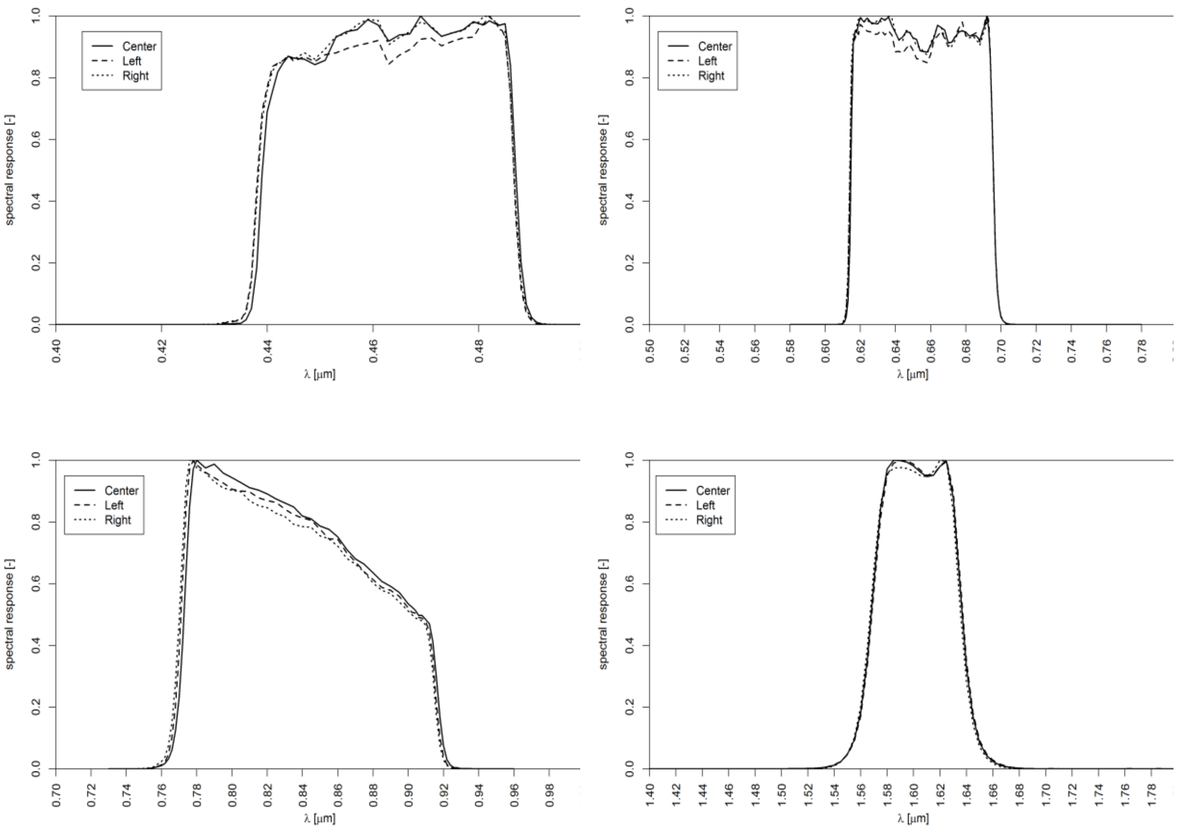


Figure 29: Spectral response functions per camera (solid=center camera, dashed=left camera, and dotted=right camera) for the BLUE (upper left), RED (upper right), NIR (lower left) and SWIR (lower right) channels.

The Spectral Response Functions are also available online at http://proba-v.vgt.vito.be/en/quality/platform-status-information/spectral-response-functions.

Appendix B: Improved cloud detection algorithm details

As shortly explained in Section 2.5.1, the major improvements of the new cloud detection algorithm are the construction of a training dataset, the introduction of high-resolution surface albedo reference maps, and the implementation of an extended set of threshold tests and similarity checks. The different algorithm steps are explained in more detail below.

Training dataset

The new cloud detection algorithm was trained on PROBA-V pixels that were manually selected. A large database of pixels was built to contain the following classes of cloud cover: 'totally cloudy', 'semi-cloudy' and 'clear'. In total, 84758 pixels were selected from four daily world composites, one per season: 21 March, 21 June, 21 September, and 21 December.

Three main land cover categories were defined for the pixel selection: 'water', 'land', and 'snow/ice'. For each category, distinct pixel classes were defined as follows:

- 'Water' category: dark water - ocean, dark water - inland, turbid water
- 'Land' category: desert, vegetation, wetland, salt, urban
- 'Snow/ice' category: snow/ice

The selected pixels are globally spread in order to take into account possible effects of high latitudes on the retrieved reflectance data. A pixel training set was randomly extracted from the database using a sampling matrix that was designed to represent all pixel classes and to preserve a good balance between 'cloudy' and 'clear' pixels.

BLUE band reference reflectances

Monthly surface albedo reference maps were obtained from average values of the two MERIS bands measuring reflectance in the BLUE region (centered at 0.413 and 0.443 µm), having the same spatial resolution as PROBA-V (300 m). These reference reflectance values were obtained for the entire MERIS mission (2002 – 2012). GlobAlbedo shortwave (0.3 – 0.7 µm) surface albedo values at 0.05° × 0.05° were used as a backup when MERIS reflectances were missing (e.g. due to lack of observations during the winter season).

Status map

In addition to the BLUE band reference values, the new cloud detection algorithm uses information related to the land cover class. Similarly to the reflectance reference values, a monthly status value was set for each pixel. Status values were provided by the Land Cover CCI product (see http://www.esa-landcover-cci.org/?q=webfm_send/75), which assigns each particular given pixel to one of the following classes: 'land', 'water', 'snow/ice', 'shadow', 'cloud', and 'unknown'. In the classification algorithm, the status map contains integer values which are assigned to a specific land cover as follows: 1 – land, 2 – water, 3 – snow/ice, 4 – other. The last class collects the three undefined classes: 'shadow', 'cloud' and 'unknown'.

Reference spectra

A set of reference spectra, based on averaged clear-sky PROBA-V observations, were composed to serve as reference for the similarity checks. These spectra are representative for specific land cover types and each of them was obtained as averages over a large number of PROBA-V pixels manually extracted from real imagery. The following reference spectra were defined:

- Clear ice
- Total cloud
- Semi-cloud over salt
- Semi-cloud over water
- Clear water
- Clear salt
- Semi-cloud over ice
- Clear turbid water
- Clear wetland
- Green water (inland lakes)
- Dark green water (ocean)
- Sea green water (river, estuaries, deltas)
- Static green water

Decision rules

In the following, “T” refers to a threshold test (per band, band ratios or amplitude differences), while “S” denotes a similarity check. The similarity between an observed spectrum *p* and a reference spectrum *R* is measured by the Spectral Angle Distance (SAD), which measures the cosine of the angle between the two vectors. Low SAD values indicate high similarity, while large SAD values exhibit low similarity. This metric was chosen as it is ideally invariant with the illumination conditions.

The possible inputs to the decision rules are:

- R_B: the Blue TOA reflectance value from the image pixel
- R_S: the SWIR TOA reflectance value from the image pixel
- p: the TOA reflectance spectrum from the image pixel in (Blue, Red, NIR, SWIR)
- $R_{B,ref}$: the TOC reflectance from the reference image at the nearest location to the image pixel

The designed set of rules contains the following checks.

Threshold tests

T1: a R_B threshold, T_B, provides an initial separation between cloudy and clear pixels, so if $R_B > T_B$, then a pixel is labelled cloudy.

T2: if the ratio $R_B / R_{B,ref}$ is lower than threshold $T_{B,ref}$, a pixel is marked as clear.

T3: if $r_B > T_{B,ref,}+T_{add}$, with T_{add} an additional threshold, then a pixel initially labelled ‘clear’ will be set to ‘cloudy’.

T4: if R_S exceeds threshold T_{S1}, a pixel initially labelled ‘clear’ will be marked as ‘cloudy’.

T5: if a ‘clear’ pixel has R_S lower than threshold $T_{S2,}$ it should be marked as ‘clear’.

T6: if a pixel is marked as cloudy, but the ratio R_B / $R_S > T_{BS}$, the pixel should be marked as ‘clear’.

Similarity checks

S1: if a pixel is marked as cloud, but its spectrum *p* is much more similar to the snow reference spectrum than to the total cloud reference spectrum, the pixel should be marked as clear.

S2: if a pixel is marked as cloud, but its spectrum *p* is similar to the snow reference spectrum, the pixel should be marked as clear.

S3: if a pixel is marked as clear, but its spectrum *p* is similar to the semi-cloudy snow reference spectrum, the pixel should be marked as cloudy.

S4: if a pixel is marked as clear, but its spectrum *p* is similar to the total cloud reference spectrum, the pixel should be marked as cloudy.

S5: if a pixel is marked as clear, but its spectrum *p* is similar to the semi-cloud reference spectrum over salt, the pixel should be marked as cloudy.

S6: if a pixel is marked as cloudy, but its spectrum *p* is similar to the clear turbid water reference spectrum, the pixel should be marked as clear.

S7: if a pixel is marked as cloudy, but its spectrum *p* is similar to the clear water reference spectrum, the pixel should be marked as clear.

S8: if a pixel is marked as cloudy, but its spectrum *p* is similar to the dark green water reference spectrum, the pixel should be marked as clear.

S9: if a pixel is marked as cloudy, but its spectrum *p* is similar to the green water reference spectrum, the pixel should be marked as clear.

S10: if a pixel is marked as cloudy, but its spectrum *p* is similar to the sea green water reference spectrum, the pixel should be marked as clear.

S11: if a pixel is marked as clear, but its spectrum *p* is similar to the semi-cloudy water reference spectrum, the pixel should be marked as cloudy.

S12: if a pixel is marked as cloudy, but its spectrum *p* is similar to the salt reference spectrum, the pixel should be marked as clear.

S13: if a pixel is marked as cloudy, but its spectrum *p* is similar to the clear wetland reference spectrum, the pixel should be marked as clear.

S14: removes a pixel from the cloud class if its spectrum is close enough to the static green water reference.

All threshold tests and similarity checks were manually tuned per land cover class to obtain the minimum overall classification error in the training dataset. Table 12 and Table 13 present the threshold values for the threshold tests and similarity checks.

The following threshold and similarity tests were applied to the different pixel statuses (land cover class):

- Pixel status 1 (land): T1, T2, T3, S1, S2, S3, S6, S7
- Pixel status 2 (water): T1, T2, T3, T4, T5, T6, S11, S7, S12, S13
- Pixel status 3 (snow/ice): T3, S2, S3, S5, S6, S13, S7
- Pixel status 4 (unknown land cover): T1, T3, S1, T5, S6, S13, S7

The global tests (no distinction on land cover) contain tests: S8, S9, S10, S14.

Table 12: Thresholds used in the final version for "T" tests

Pixel status	T1	T2	T3	T4	T5	T6
1	0.24	1.5	0.13	-	-	-
2	0.174	1.9	0.13	0.46	0.03	7.4
3	-	-	0	-	-	-
4	0.14	-	0.09	-	-	-

Table 13: Thresholds used in the final version for the "S" tests

<table>
<tr><th>Pixel status</th><th>S1</th><th>S2</th><th>S3</th><th>S4</th><th>S5</th><th>S6</th><th>S7</th><th>S8</th><th>S9</th><th>S10</th><th>S11</th><th>S12</th><th>S13</th><th>S14</th></tr>
<tr><td>1</td><td>0.11</td><td>0.05</td><td>0.06</td><td>-</td><td>-</td><td>0.14</td><td>0.1</td><td rowspan="4">0.12</td><td rowspan="4">0.14</td><td rowspan="4">0.14</td><td>-</td><td>-</td><td>-</td><td rowspan="4">0.06</td></tr>
<tr><td>2</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>0.07</td><td>0.09</td><td>0.01</td><td>0.23</td></tr>
<tr><td>3</td><td>-</td><td>0.3</td><td>0.07</td><td>-</td><td>0.13</td><td>0.06</td><td>0.58</td><td>-</td><td>-</td><td>0.23</td></tr>
<tr><td>4</td><td>0.09</td><td>-</td><td>-</td><td>-</td><td>-</td><td>0.06</td><td>0.02</td><td>-</td><td>-</td><td>0.23</td></tr>
</table>

Appendix C: Collection 1 ICP file update details

This Appendix contains more details on the ICP file update for the Collection 1 reprocessing as presented in Section 2.5.2, with the enumeration applied analogously. New ICP files were created per month for the total reprocessing period.

1. Inter-camera adjustments to VNIR absolute calibration coefficients

In order to improve the consistency between the three PROBA-V cameras, some adjustments to the absolute calibration coefficients of the VNIR strips were made after the commissioning phase, to correct for a band-dependent camera-to-camera bias. On 26 June 2014 a 1.8% radiance reduction to the BLUE CENTER strip was applied and a 1.2% radiance increase was applied to the BLUE RIGHT strip. On 23 September 2014 a 2.1% reduction to the RED CENTER radiance, a 1.1% increase to the BLUE LEFT radiance, and a 1.3% increase to the NIR LEFT radiance was applied.
Finally, on 25 October 2014 the RIGHT NIR radiance was increased by 1%.
The reprocessing includes the inter-camera bias adjustment from the mission start. In Figure 30, the impact on the TOA reflectance is illustrated. Please note that there are no changes to the following VNIR strips: RED LEFT, RED RIGHT, and NIR CENTER.

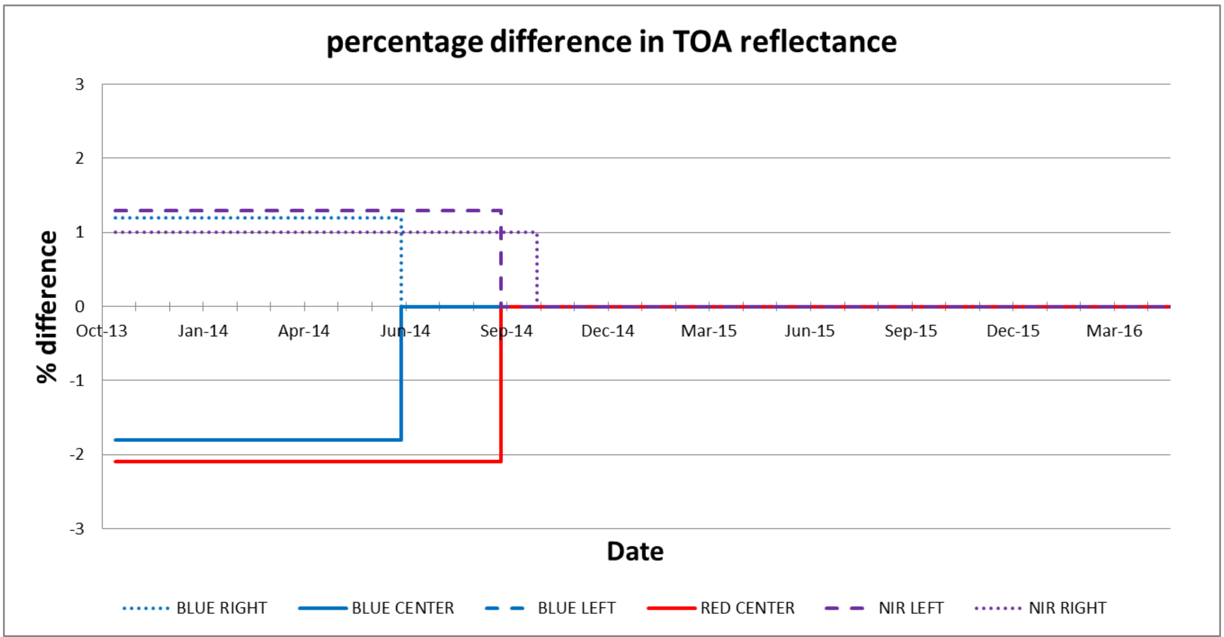


Figure 30: Relative difference [%] in TOA reflectance between reprocessed and old archive as a function of the acquisition date.

2. Application of a degradation model to the SWIR absolute calibration coefficients

A linear trending model is fitted to the Optical Sensor Calibration with Simulated Radiance (OSCAR) desert vicarious calibration results obtained for the nine different SWIR strips. Using this degradation model in the reprocessing, the step-wise adjustment of the SWIR absolute calibration coefficients is replaced by a smooth and gradual correction, as illustrated in Figure 31.

3. Improvement multi-angular calibration SWIR strips CENTER camera

In order to better characterize and to correct for non-uniformities within and between detectors, a 90° yaw maneuver has been performed with PROBA-V over the Niger-1 desert site on 11 April 2016. With this 90° yaw configuration the detector array runs parallel to the direction of motion and therefore an area on the ground is subsequently viewed by the different pixels of the strip. From the acquired yaw data improved low frequency multi-angular calibration coefficients have been derived. Figure 32 presents the changes to the equalization coefficients of the different CENTER camera SWIR strips.

Figure 32 illustrates the impact applying of the improved equalisation coefficients. All other strips are excluded, as no final profiles could be derived from the yaw data.

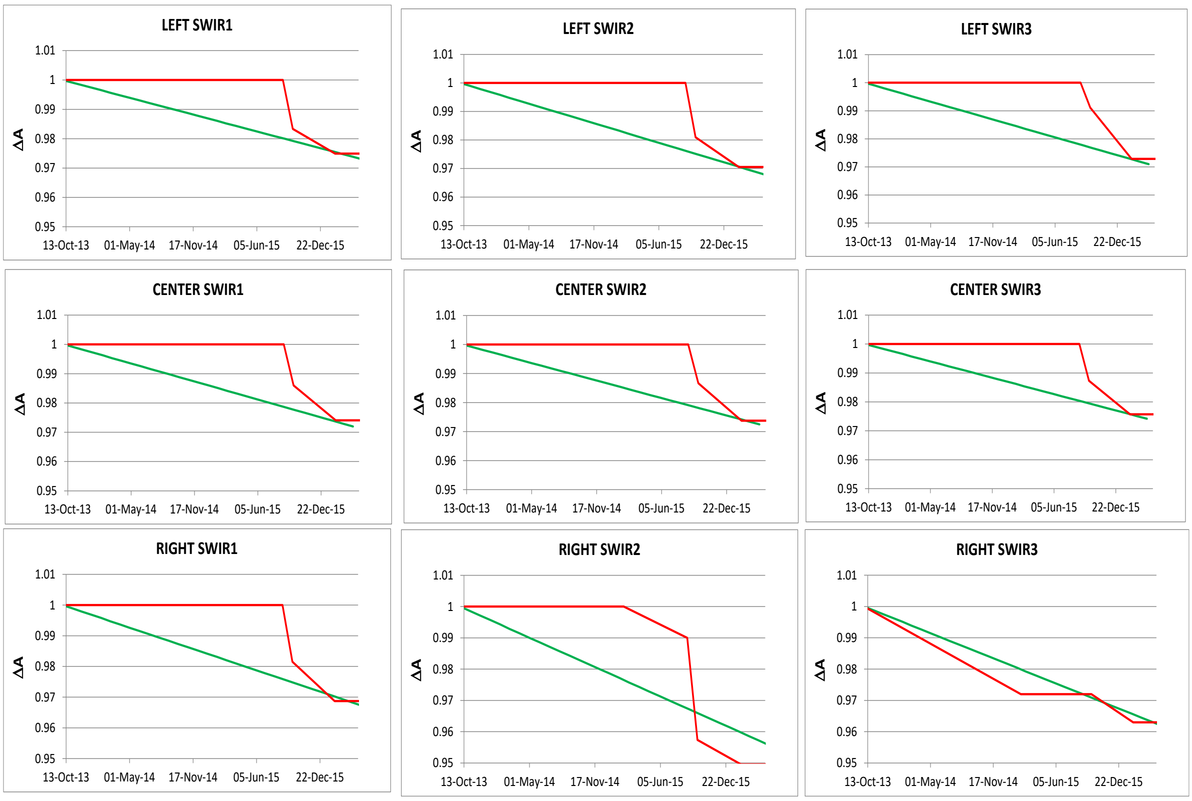


Figure 31: Evolution of the absolute calibration coefficient over time (red line: old, green line: reprocessing). A value lower than 1 results in a TOA reflectance increase.

4. Dark current

ICP files for reprocessing of data acquired before 2015: Before 2015 the operational dark current calibration was based on dark current acquisitions with a very long integration time of 3 s. It was noticed that for some SWIR pixels with a very high dark current detector saturation and/or non-linearity effects occurred for this long integration time. In the Collection 1 reprocessing, dark current values of saturated pixels were replaced with the value retrieved from the closest dark current observation obtained with a shorter integration time. Furthermore, the ICP files were corrected for a bug that caused the assignment of the dark current to the wrong pixel ID in the generation of the final xml-formatted ICP file.

For all ICP files: dark current values are based on dark current acquisitions of the applicable month, while in the old NRT processing dark current values were based on acquisitions of the previous month.

5. Bad pixels

A pixel is declared bad at the start of a month aligned with the starting date of an ICP update.

Figure 32: Changes to the equalisation over the field of view. A value lower than 1 results in an increase in the TOA reflectance while a value higher than 1 gives a TOA reflectance decrease.

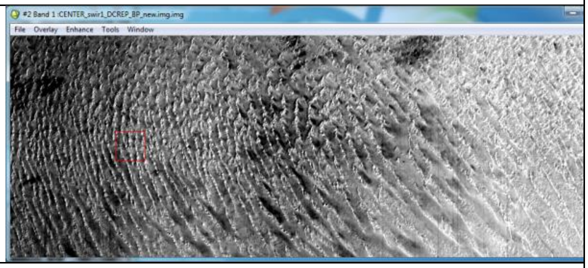
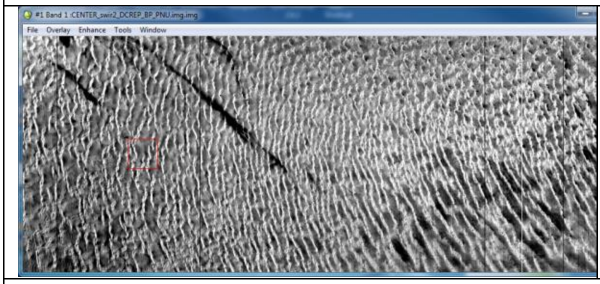
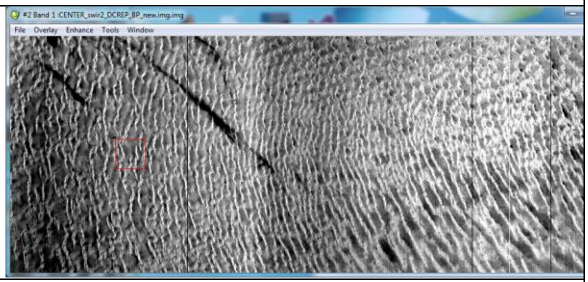
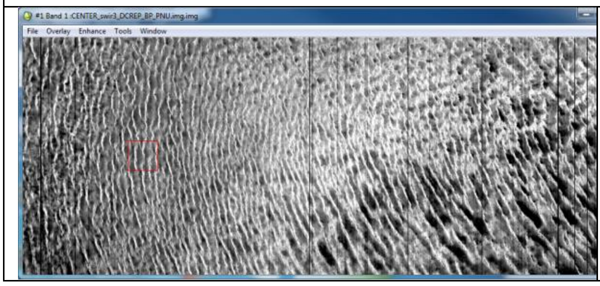
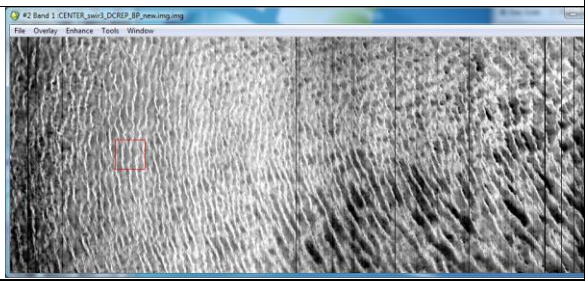
Reprocessed	Old
CENTER SWIR1	
	
CENTER SWIR2	
	
CENTER SWIR3	
	

Figure 33: Illustration of the impact of improved equalisation parameters on the scene uniformity. Vertical lines indicate bad pixels.

Appendix D: Collection 0 cloud detection algorithm description

The Collection 0 cloud detection algorithm was a modified version of the method applied to the SPOT-VGT BLUE and SWIR observations (Lissens et al., 2000). Using these band reflectances, two separate cloud masks are created. A 3 × 3-pixel search mask is applied to determine the matching SWIR pixel for the BLUE band cloud mask, and the matching BLUE pixel for the SWIR band cloud mask (see Figure 34).The final cloud mask is a merge of these two masking results. Compared to the SPOT-VGT cloud mask, some modifications were necessary, because the assumption that clouds are observed at the same position in both the BLUE and SWIR bands is no longer valid for PROBA-V, due to the observation time difference. This is further explained below.

As already indicated in Section 1, the PROBA-V instrument design is such that the NIR observes a cloud first, followed by the RED, BLUE, and SWIR bands. The time difference between the NIR and SWIR cloud observations is 12 s. As a result, the NIR and SWIR bands will map clouds onto different positions in the along-track direction, with differences up to ~700 m for clouds at 10 km altitude. Other effects of the observation time difference include viewing angle differences and horizontal cloud shifts. The maximum shift resulting from the latter two effects will not exceed one 300 m pixel along-track and one pixel cross-track on either side.

The cloud detection algorithm accounts for the different observation times as follows.
For the cloud detection based on the BLUE band reflectance, it is checked whether the observed value exceeds the BLUE band reflectance threshold of 0.2465. In addition, it is checked whether the maximum SWIR reflectance value in a 3x3 pixel box (i.e., 1x1 km) above the BLUE pixel in the image exceeds the SWIR band threshold, as depicted in the upper panels in Figure 34. If both conditions are satisfied, the BLUE pixel is classified as cloudy. Note that this 3x3-pixel SWIR reflectance test only needs to be applied to the 300 m cloud mask.
Because pixels observed in NIR are mostly observed in front of the BLUE pixel in the image, the pixel below the BLUE pixel is also categorized as cloudy.

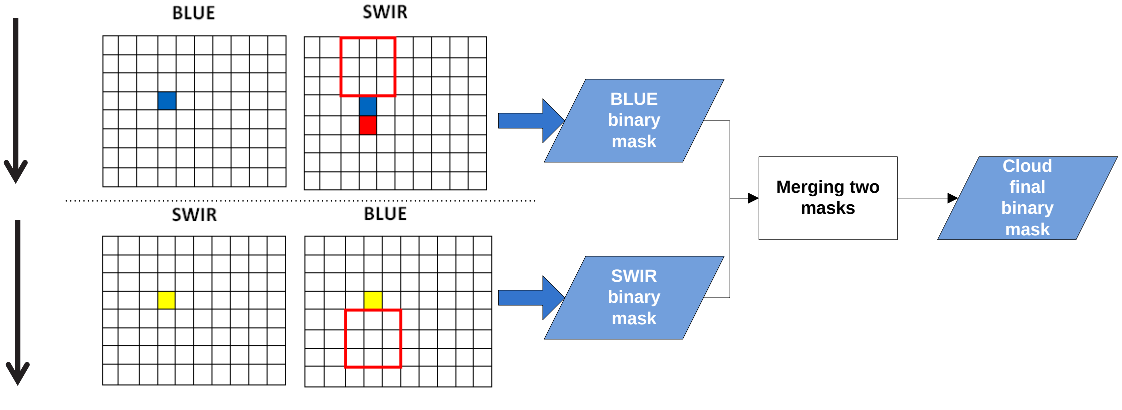


Figure 34: The Collection 0 cloud detection process for the BLUE, SWIR, and final cloud masks. The red pixel denotes the NIR observation. The satellite along-track flight direction is indicated by the black arrows. The 3x3-pixel SWIR reflectance test is only applied for the 300 m cloud mask.

A similar procedure was applied for the SWIR-based cloud mask, but then with an additional forward 3×3 BLUE pixel window (see the lower panel of Figure 34). The reflectance threshold value for the SWIR band to label a pixel as ‘cloudy’ is 0.09. The final cloud mask is obtained through merging the BLUE and SWIR masks, with values of 0 and 1 indicating ‘clear’ and ‘cloudy’, respectively.

Appendix E1: Detailed Level 1C Product file description

Below a detailed description of the Segment Product (LEVEL 1C) files is given. Reference is made to Figure 18, which presents the dataset structure of the file. Note that for the entire Collection 1 CF compliant metadata were added. More information, as well as an overview of the metadata fields is given in Section 4.5 and Table 8, respectively.

Table 14: HDF5 structure of LEVEL 1C product file.

Type	Name	Description	Data type
HDF5 Group	LEVEL1A	HDF5 group containing the Level1A data and metadata. The structure and content of this group is elaborated in Table 15.	See Table 15
HDF5 Group	LEVEL1B	HDF5 group containing the Level1B data and metadata. The structure and content of this group is elaborated in Table 16.	See Table 16
HDF5 Group	LEVEL 1C	HDF5 group containing the LEVEL 1C data and metadata. The structure and content of this group is elaborated in Table 17.	See Table 17
HDF5 Attribute	ACQUISITION_STATION	Identifier for the data reception station (i.e., Kiruna or Fairbanks).	String
HDF5 Attribute	BOTTOM_LEFT_LATITUDE	The latitude of the bottom-left point of the bounding box [º].	32-bit floating-point
HDF5 Attribute	BOTTOM_LEFT_LONGITUDE	The longitude of the bottom-left point of the bounding box [º].	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_LATITUDE	The latitude of the bottom-right point of the bounding box [º].	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_LONGITUDE	The longitude of the bottom-right point of the bounding box [º].	32-bit floating-point
HDF5 Attribute	CAMERA	Camera identifier. Possible values are: “LEFT” for the left camera (camera 1), “CENTER” for the center camera (camera 2), “RIGHT” for the right camera (camera 3)	String
HDF5 Attribute	DATELINE_CROSSING	Flag indicating whether or not the segment is crossing the dateline (180° border).	String
HDF5 Attribute	DAY_NIGHT_FLAG	Indicating whether or not the segment is a day segment or taken at night.	String
HDF5 Attribute	DEFECT_PIXELMAP_ID	Field identifying the defect pixel map.	32-bit integer
HDF5 Attribute	DESCRIPTION	Short description of the file content, i.e. “PROBA-V LEVEL 1C product”.	String
HDF5 Attribute	INSTRUMENT	Short name for the instrument, i.e. VEGETATION.	String
HDF5 Attribute	NORTHPOLE_INDICATOR	Flag indicating whether or not the segment is covering the north pole.	String

Type	Name	Description	Data type
HDF5 Attribute	NUMBER_OF_STRIPS	The number of strips this product contains. This value is typically set to 6.	32-bit integer
HDF5 Attribute	OBSERVATION_END_DATE	The observation end date (UTC) of the segment. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_END_TIME	The observation end time (UTC) of the segment. The format is: hh:mm:ss.μμμμμμ.	String
HDF5 Attribute	OBSERVATION_START_DATE	The observation start date (UTC) of the segment. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_START_TIME	The observation start time (UTC) of the segment. The format is: hh:mm:ss.μμμμμμ.	String
HDF5 Attribute	OVERPASS_NUMBER	The overpass number.	32-bit integer
HDF5 Attribute	PLATFORM	Short name for the platform and its serial number, i.e. PROBA-1.	String
HDF5 Attribute	PROCESSING_DATE	The date the product was generated. The format is: YYYY-MM-DD.	String
HDF5 Attribute	PROCESSING_TIME	The time the product was generated. The format is: hh:mm:ss.μμμμμμ.	String
HDF5 Attribute	PRODUCT_REFERENCE	A unique textual reference to the product (type: string). This string has following syntax: RawSegment_PROBAV#<CAMERA>_<YYYYMMDD><hhmmss>_LEVEL1C_V<VERSION> Where: <CAMERA>: identifier for the camera (1, 2 or 3) <YYYYMMDD>: the observation start date <hhmmss>: the start observation start time <VERSION>: the version identifier (three digits)	String
HDF5 Attribute	ROI_IDENTIFIER	Identifier for the Region Of Interest. If the LEVEL 1C product is a full swath product, it contains the value "FULL_SWATH".	String
HDF5 Attribute	SOUTHPOLE_INDICATOR	Flag indicating whether or not the segment is covering the south pole.	String
HDF5 Attribute	TOP_LEFT_LATITUDE	The latitude of the top-left point of the bounding box [°].	32-bit floating-point
HDF5 Attribute	TOP_LEFT_LONGITUDE	The longitude of the top-left point of the bounding box [°].	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_LATITUDE	The latitude of the top-right point of the bounding box [°].	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_LONGITUDE	The longitude of the top-right point of the bounding box [°].	32-bit floating-point

Table 15: HDF5 structure of LEVEL1A Group.

Type	Name	Description
HDF5 Group	PLATFORM	HDF5 group containing the platform data and ancillary data that is applicable for this segment. The content and structure of this group is elaborated in Table 16.
HDF5 Group	BLUE	HDF5 group containing the instrument data and metadata for the BLUE strip.The structure and content of this group is elaborated in Table 17.
HDF5 Group	NIR	HDF5 group containing the instrument data and metadata for the NIR strip. The structure and content of this group is elaborated in Table 17.
HDF5 Group	RED	HDF5 group containing the instrument data and metadata for the RED strip. The structure and content of this group is elaborated in Table 17.
HDF5 Group	SWIR1	HDF5 group containing the instrument data and metadata for the SWIR1 strip. The structure and content of this group is elaborated in Table 17.
HDF5 Group	SWIR2	HDF5 group containing the instrument data and metadata for the SWIR2 strip. The structure and content of this group is elaborated in Table 17.
HDF5 Group	SWIR3	HDF5 group containing the instrument data and metadata for the SWIR3 strip. The structure and content of this group is elaborated in Table 17.

Table 16: HDF5 structure of PLATFORM Group

Type	Name	Description	Data type
HDF5 Table	ATT	Table containing the attitude-related data with a frequency of 4 Hz. The table consists of the following fields: • MJD: the Modified Julian Date in TAI (Temps Atomique International). • QW: the first quaternion (BodyFixed frame (BOF) to Celestial frame (CEL). • QX: the second quaternion (BodyFixed frame (BOF) to Celestial frame. • QY: the third quaternion (BodyFixed frame (BOF) to Celestial frame. • QZ: the fourth quaternion (BodyFixed frame (BOF) to Celestial frame. The table contains following attributes: • FREQUENCY: the frequency [Hz] at which the ATT data is generated. • VERSION: the version number. Every field of the table contains following attributes: • DESCRIPTION: short description of the field content • NAME: the name of the field • UNITS: the units or “-“ in case no units are available.	64-bit floating point 64-bit floating point 64-bit floating point 64-bit floating point 64-bit floating point 32-bit floating point String String String string
HDF5 Table	OBET_ GPS	Table containing the time correlation data between on Board Elapsed Time (OBET) and GPS with a frequency of 1 Hz. The table consists of following fields: • OBET: the OBET time [s] • GPS_WEEK: the GPS week	64-bit floating point 64-bit floating point 16-bit integer

		• GPS_SECONDS: the GPS second time [s] The table contains the following attributes: • FREQUENCY: the frequency [Hz] at which the time correlation related data is generated. • VERSION: the version number. Every field of the table contains the following attributes: • DESCRIPTION: short description of the field content • NAME: the name of the field • UNITS: the units or "-" in case no units are available.	64-bit floating point 32-bit floating point String String String String
HDF5 Table	OBET_VI	Table containing the time correlation data between OBET and VI (Vegetation Instrument) with a frequency of 1 Hz. The table consists of following fields: • OBET: the OBET time [s]. • VI: the VI time [s]. The table contains following attributes: • FREQUENCY: the frequency [Hz] at which the time correlation related data is generated. Every field of the table contains following attributes: • DESCRIPTION: short description of the field content . • NAME: the name of the field. • UNITS: the units or "-" in case no units are available.	64-bit floating point 64-bit floating point 32-bit floating point String String string
HDF5 Table	PRM	Table including any housekeeping telemetry generated by the Vegetation Instrument. The data has a frequency of 1 Hz. The table consists of following fields: • MJD: the Modified Julian Date in TAI [d] (Temps Atomique International, data type: 64-bit floating-point). • TIME_OUT_ECLIPSE: the time since out of eclipse (delta) [s]. The special value 0 means in eclipse and the value -1 means NO VALUE. • VI_1 : VI Temperature TMA of left spectral imager. • VI_2 : VI Temperature TMA of central spectral imager. • VI_3 : VI Temperature TMA of right spectral. • VI_4 : VI parameter 4 TBD. • VI_5: VI parameter 5 TBD. • VI_6: VI parameter 6 TBD. • VI_7: VI parameter 7 Temperature: Optical bench near star tracker. • VI_8: VI parameter 8 Temperature: Left SWIR detector. • VI_9: VI parameter 9 Temperature: Central SWIR detector. • VI_10: VI parameter 10 Temperature: Right SWIR detector. • VI_11: VI parameter 11 Temperature: Left VNIR detector. • VI_12: VI parameter 12 Temperature: Central VNIR detector. • VI_13: VI parameter 13 Temperature: Right VNIR detector	32-bit unsigned integer (through SPARE_8)

		• VI_14: VI parameter 14 Temperature: Left flexure. • VI_15: VI parameter 15 Temperature: Central flexure • VI_16:VI parameter 16 Temperature: Right flexure • SPARE_1: spare parameter TBD. • SPARE_2: spare parameter TBD. • SPARE_3: spare parameter TBD. • SPARE_4: spare parameter TBD. • SPARE_5: spare parameter TBD. • SPARE_6: spare parameter TBD. • SPARE_7: spare parameter TBD. • SPARE_8: spare parameter TBD. • SPARE_9: spare parameter TBD. • SPARE_10: spare parameter TBD. • SPARE_11: spare parameter TBD. • SPARE_12: spare parameter TBD. • SPARE_13: spare parameter TBD. • SPARE_14: spare parameter TBD. • SPARE_15: spare parameter TBD. • SPARE_16: spare parameter TBD. The table contains following attributes: • FREQUENCY: the frequency [Hz]. • PACKET_EDITION: the packet edition. • VERSION: the current version. Every field of the table contains following attributes: • DESCRIPTION: short description of the field content . • NAME: the name of the field. • UNITS: the units.	64-bit unsigned integer (through SPARE_16) 32-bit unsigned integer 32-bit unsigned integer String String String string
HDF5 Table	PV	Table containing the position and velocity related data with a frequency of 1 Hz. The table consists of the following fields: • MJD: the Modified Julian Date in TAI (Temps Atomique International). • PX: the position in the X direction in the Terrestrial frame (TER) [m]. • PY: the position in the Y direction in the Terrestrial frame (TER) [m]. • PZ: the position in the Z direction in the Terrestrial frame (TER) [m]. • VX: the velocity in the X direction in the Terrestrial frame (TER) [m s^{-1}]. • VY: the velocity in the Y direction in the Terrestrial frame (TER) [m s^{-1}]. • VZ: the velocity in the Z direction in the Terrestrial frame (TER) [m s^{-1}]. The table contains following attributes: • FREQUENCY: the frequency [Hz] at which the PV data is generated. • VERSION: the version number. Every field of the table contains following attributes:	64-bit floating point (through VZ) 32-bit floating point String

		• DESCRIPTION: short description of the field content. • NAME: the name of the field. • UNITS: the units.	String String string

Table 17: HDF5 structure of LEVEL1A STRIP (BLUE, RED, NIR, SWIR1, SWIR2, and SWIR3) Groups.

Type	Name	Description	Data type
HDF5 Table	BLOCKDATA	Table containing the block related data. The table consists of following fields: • MJD: the Modified Julian Date in TAI (Temps Atomique International), [d]. • INTEGRATION_TIME: the integration time [s] • TEMPERATURE: the temperature [° C] () • DARK_PIXEL_1: the dark pixel value for pixel 3 • DARK_PIXEL_2: the dark pixel value for pixel 4 • DARK_PIXEL_3: the dark pixel value for pixel 5997 • DARK_PIXEL_4: the dark pixel value for pixel 5998 Each field contains following attributes: • DESCRIPTION: short description of the field content • NAME: the name of the field • UNITS: the units **Note:** there are no BLOCKDATA dataset for the SWIR strips (SWIR1, SWIR2, SWIR3)	 64-bit floating-point 32-bit floating point 32-bit floating-point 16-bit integer 16-bit integer 16-bit integer 16-bit integer string string string

HDF5 Table	LINEDATA	Table containing the line related data. The table consists of the following fields:
		• MJD: the Modified Julian Date in TAI (Temps Atomique International) [d] associated with the centre of the integration period.
		• LINE_QUALITY: line flag indicating whether a line is :
		  ▪ good (flag = 0)
		  ▪ bad (flag = 1)
		  ▪ missing (flag = 2)
		• INTEGRATION_TIME: the integration time [s]
		• TIME_OUT_ECLIPSE: the time since out of eclipse (delta) [s]. The special value "0" means in eclipse and the value "-9999" means NO_DATA.
		• TEMPERATURE: the temperature [° C] (NO_DATA = -9999)
			
			64-bit floating-point
			
			32-bit integer
			
			32-bit floating-point
			32-bit integer
			
			32-bit floating-point

Type	Name	Description	Data type
		Further, every field of the table contains following attributes: • DESCRIPTION: short description of the field content (data type: string) • NAME: the name of the field NO_DATA: the "no data" value. This attribute is optional in case no "no data" value is applicable for the field. • UNITS: the used units	String String NO_DATA: 32-bit integer or float string
HDF5 Dataset	DN	Dataset containing the digital numbers (extracted from the raw data). Table 18 lists the metadata items attached to this dataset.	16-bit integer
HDF5 Attribute	COMPRESSION_RATIO	The used compression ratio.	32-bit floating-point
HDF5 Attribute	DETECTOR	Identifier for the detector. Possible values are: VNIR, SWIR	String
HDF5 Attribute	GAIN_FACTOR	The gain factor.	32-bit floating-point
HDF5 Attribute	LINE_END	The end line of the bottom-right pixel value in the image.	32-bit integer
HDF5 Attribute	LINE_START	The start line of the top-left pixel value (0, 0) in the image. In case the image contains the full swatch, this value is set to 0.	32-bit integer
HDF5 Attribute	OBSERVATION_END_DATE	The observation end date (UTC), i.e. the date of the last line of the strip. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_END_TIME	The observation end time (UTC), i.e., the time of the last line of the strip. The format is: hh:mm:ss.μμμμμμ.	String
HDF5 Attribute	OBSERVATION_START_DATE	The observation start date (UTC), i.e., the date of the first line of the strip. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_START_TIME	The observation start time (UTC), i.e., the time of the first line of the strip. The format is: hh:mm:ss.μμμμμμ.	String
HDF5 Attribute	PIXEL_END	The end index of the bottom-right pixel value in the image. In case the image contains the full swath, this value is set to: 5199 for the VNIR strips 1023 for the SWIR strips	32-bit integer
HDF5 Attribute	PIXEL_START	The start index of the top-left pixel value (0, 0) in the image. In case the image contains the full swatch, this value is set to 0.	32-bit integer
HDF5 Attribute	STRIP	Identifier for the strip. Possible values are: BLUE, RED, NIR, SWIR1 (left SWIR strip), SWIR2 (center SWIR strip), SWIR3 (right SWIR strip)	String

Table 18: HDF5 metadata items for DN datasets.

Type	Name	Description	Data type
HDF5 Attribute	DESCRIPTION	Description of the DN dataset name.	String

Type	Name	Description	Data type
HDF5 Attribute	MAPPING	The mapping information, consisting of following values: • **<proj_id>**: the projection ID (e.g. "Geographic Lat/Lon") • **<x_m>**: A value indicating whether the map X coordinates refer to the top-left corner (0.0) or center (0.5) of the pixel. • **<y_m>**: A value indicating whether the map Y coordinates refer to the top-left corner (0.0) or center (0.5) of the pixel. • **<x_start>**: the X coordinate of the upper-left pixel. • **<y_start>**: the Y coordinate of the upper-left pixel. • **<x_res>**: the spatial resolution in the X direction. • **<y_res>**: the spatial resolution in the Y direction. • **<datum>**: the projection's datum (in case of unprojected image, the value is '-'). • **<units>**: the projection's unit (in case of an unprojected image, the value is '-'). Note that his is an optional attribute. If not provided, it is assumed that no geographical information is attached to the dataset.	String
HDF5 Attribute	NO_DATA	Indicates the "no data" value of the dataset. This value has been set to "-1".	32-bit floating-point
HDF5 Attribute	OFFSET	OFFSET values (see SCALE attribute)	32-bit floating-point
HDF5 Attribute	SCALE	The coding information, indicating that the pixels have been scaled with a given scale and offset according to following formula: DN = OFFSET + (PV * SCALE) • The scale factor • The offset factor The physical value is determined as PV = (DN-OFFSET) / SCALE.	32-bit floating-point

Table 19: HDF5 structure of LEVEL1B group.

Type	Name	Description	Data type
HDF5 Table	CONTOUR	Table containing the contour points of the segment. The table contains two fields: • The longitude values of the segment's contour. • The latitude values of the segment's contour. This contour is the contour that encloses all the strip contours.	32-bit floating point
HDF5 Group	BLUE	HDF5 group containing the instrument data and metadata for the BLUE strip. The structure and content of this group is elaborated in Table 20.	See Table 20
HDF5 Group	NIR	HDF5 group containing the instrument data and metadata for the NIR strip. The structure and content of this group is elaborated in Table 20.	See Table 20

Type	Name	Description	Data type
HDF5 Group	RED	HDF5 group containing the instrument data and metadata for the RED strip. The structure and content of this group is elaborated in Table 20.	See Table 20
HDF5 Group	SWIR1	HDF5 group containing the instrument data and metadata for the SWIR1 strip. The structure and content of this group is elaborated in Table 20.	See Table 20
HDF5 Group	SWIR2	HDF5 group containing the instrument data and metadata for the SWIR2 strip. The structure and content of this group is elaborated in Table 20.	See Table 20
HDF5 Group	SWIR3	HDF5 group containing the instrument data and metadata for the SWIR3 strip. The structure and content of this group is elaborated in Table 20.	See Table 20
HDF5 Attribute	ICP_REFERENCE	Reference to the used geometric ICP file. This string has the following syntax: PROBAV_ICP_GEOMETRIC#{LEFT|RIGHT|CENTER}_<StartDate>_V<Revision>	String
HDF5 Attribute	LEAPSECONDS	Leap second =TAI-UTC [s].	32-bit floating point
HDF5 Attribute	POLARMOTION_ DELTA_UT1	The difference between UT1 and UTC (UT1-UTC), [s].	32-bit floating point
HDF5 Attribute	POLARMOTION_X	The X position of the Celestial Intermediate Pole (CIP) and the Celestial/Terrestrial Ephemeris Origins (CEO, TEO) in the Geocentric Celestial Reference System (GCRS) and International Terrestrial Reference System (ITRS) [sec].	32-bit floating point
HDF5 Attribute	POLARMOTION_Y	The Y position of the Celestial Intermediate Pole (CIP) and the Celestial/Terrestrial Ephemeris Origins (CEO, TEO) in the Geocentric Celestial Reference System (GCRS) and International Terrestrial Reference System (ITRS) [sec].	32-bit floating point
HDF5 Attribute	PROCESSINGINFO_ GEOMODELLING	Reference to the used geo-modelling tool, e.g. “GEOMODELLING V1.0”.	String
HDF5 Attribute	SUN_BETA_ANGLE	The sun beta angle [°].	32-bit floating point

Table 20: HDF5 structure of LEVEL1B STRIP (BLUE, NIR, RED, SWIR1, SWIR2, and SWIR3) Groups.

Type	Name	Description	Data type
HDF5 Table	CONTOUR	Table containing the contour points of the strip. The table contains: • Longitude, the longitude values of the segment's contour • Latitude, the latitude values of the segment's contour	32-bit floating point
HDF5 Dataset	LN1	Dataset containing the longitude at 0 m altitude. Table 21 lists the metadata items specific for this dataset.	32-bit floating-point
HDF5 Dataset	LN2	Dataset containing the longitude at 5000 m altitude. Table 21 lists the metadata items specific for this dataset.	32-bit floating-point
HDF5 Dataset	LT1	Dataset containing the latitude at 0 m altitude. Table 21 lists the metadata items specific for this dataset.	32-bit floating-point
HDF5 Dataset	LT2	Dataset containing the latitude at 5000 m altitude. Table 21 lists the metadata items specific for this dataset.	32-bit floating-point
HDF5 Dataset	SAA	Dataset containing the solar azimuth angles. Table 21 lists the metadata items specific for this dataset.	8-bit unsigned integer
HDF5 Dataset	SZA	Dataset containing the solar SZA. Table 21 lists the metadata items specific for this dataset.	8-bit unsigned integer
HDF5 Dataset	VAA	Dataset containing the viewing azimuth angles. Table 21 lists the metadata items specific for this dataset.	8-bit unsigned integer
HDF5 Dataset	VZA	Dataset containing the VZA. Table 21 lists the metadata items specific for this dataset.	8-bit unsigned integer
HDF5 Attribute	BOTTOM_LEFT_CORNER_LATITUDE	The latitude of the bottom-left corner point of the strip [°].	32-bit floating-point
HDF5 Attribute	BOTTOM_LEFT_CORNER_LONGITUDE	The longitude of the bottom-left corner point of the strip [°].	32-bit floating-point
HDF5 Attribute	BOTTOM_LEFT_CORNER_X	The X position of the bottom-left corner point of the strip.	32-bit floating-point
HDF5 Attribute	BOTTOM_LEFT_CORNER_Y	The Y position of the bottom-left corner point of the strip.	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_CORNER_LATITUDE	The latitude of the bottom-right corner point of the strip [°].	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_CORNER_LONGITUDE	The longitude of the bottom-right corner point of the strip [°].	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_CORNER_X	The X position of the bottom-right corner point of the strip.	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_CORNER_Y	The Y position of the bottom-right corner point of the strip.	32-bit floating-point
HDF5 Attribute	CENTER_LATITUDE	The latitude of the center point of the strip [°].	32-bit floating-point
HDF5 Attribute	CENTER_LONGITUDE	The longitude of the center point of the strip [°].	32-bit floating-point
HDF5 Attribute	CENTER_X	The X position of the center point of the strip.	32-bit floating-point
HDF5 Attribute	CENTER_Y	The Y position of the center point of the strip.	32-bit floating-point
HDF5 Attribute	DETECTOR	Identifier for the detector. Possible values are: VNIR, SWIR	String
HDF5 Attribute	OBSERVATION_END_DATE	The observation end date (UTC), i.e., the date of the last line of the strip. The format is: YYYY-MM-DD.	String

Type	Name	Description	Data type
HDF5 Attribute	OBSERVATION_ END_TIME	The observation end time (UTC), i.e. the time of the last line of the strip. The format is: hh:mm:ss.µµµµµµ.	String
HDF5 Attribute	OBSERVATION_ START_DATE	The observation start date (UTC), i.e., the date of the first line of the strip. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_ START_TIME	The observation start time (UTC), i.e., the time of the first line of the strip. The format is: hh:mm:ss.µµµµµµ.	String
HDF5 Attribute	STRIP	Identifier for the strip. Possible values are: BLUE, RED, NIR, SWIR1 (left SWIR strip), SWIR2 (center SWIR strip), SWIR3 (right SWIR strip).	String
HDF5 Attribute	TOP_LEFT_ CORNER_ LATITUDE	The latitude of the top-left corner point of the strip [°].	32-bit floating-point
HDF5 Attribute	TOP_LEFT_ CORNER_ LONGITUDE	The longitude of the top-left corner point of the strip [°].	32-bit floating-point
HDF5 Attribute	TOP_LEFT_ CORNER_X	The X position of the top-left corner point of the strip.	32-bit floating-point
HDF5 Attribute	TOP_LEFT_ CORNER_Y	The Y position of the top-left corner point of the strip.	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_ CORNER_ LATITUDE	The latitude of the top -right corner point of the strip [°].	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_ CORNER_ LONGITUDE	The longitude of the top-right corner point of the strip [°].	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_ CORNER_X	The X position of the top-right corner point of the strip.	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_ CORNER_Y	The Y position of the top-right corner point of the strip.	32-bit floating-point

Table 21: HDF5 metadata items for L1B datasets.

Type	Name	Description	Data type
HDF5 Attribute	DESCRIPTION	Description of the dataset.	String
HDF5 Attribute	MAPPING	The mapping information, consisting of following values: • **<proj_id>**: the projection ID (e.g. “Geographic Lat/Lon”) • **<x_m>**: A value indicating whether the map X coordinates refer to the top-left corner (0.0) or center (0.5) of the pixel. • **<y_m>**: A value indicating whether the map Y coordinates refer to the top-left corner (0.0) or center (0.5) of the pixel. • **<x_start>**: the X coordinate of the upper-left pixel. • **<y_start>**: the Y coordinate of the upper-left pixel. • **<x_res>**: the spatial resolution in the X direction. • **<y_res>**: the spatial resolution in the Y direction. • **<datum>**: the projection‘s datum (in case of unprojected image, the value is ‘-’). • **<units>**: the projection ‘s unit (in case of unprojected image, the value is ‘-’). Note that his is an optional attribute. If not provided, it is assumed that no geographical information is attached to the dataset.	String

Type	Name	Description	Data type
HDF5 Attribute	NO_DATA	Indicates the “no data” value of the dataset. Typically, this value is set to “-9999”. This attribute is optional, meaning that in case the attribute is not present, no “no data” value is applicable for the dataset.	String
HDF5 Attribute	OFFSET	OFFSET values (see SCALE attribute)	32-bit floating-point
HDF5 Attribute	SCALE	The coding information, indicating that the pixels have been scaled with a given scale and offset according to following formula: DN = OFFSET + (PV * SCALE) • The scale factor • The offset factor The physical value is determined as PV = (DN - OFFSET)/SCALE.	32-bit floating-point
HDF5 Attribute	UNIT	Unit type of the dataset (if not applicable the values ‘-‘ is used)	String

Table 22: HDF5 structure of LEVEL 1C group.

Type	Name	Description	Data type
HDF5 Group	BLUE	HDF5 group containing the instrument data and metadata for the BLUE strip. The structure and content of this group is elaborated in Table 23.	See Table 22.
HDF5 Group	NIR	HDF5 group containing the instrument data and metadata for the NIR strip. The structure and content of this group is elaborated in Table 23.	See Table 22.
HDF5 Group	RED	HDF5 group containing the instrument data and metadata for the RED strip. The structure and content of this group is elaborated in Table 23.	See Table 22.
HDF5 Group	SWIR1	HDF5 group containing the instrument data and metadata for the SWIR1 strip. The structure and content of this group is elaborated in Table 23.	See Table 22.
HDF5 Group	SWIR2	HDF5 group containing the instrument data and metadata for the SWIR2 strip. The structure and content of this group is elaborated in Table 23.	See Table 22.

HDF5 Group	SWIR3	HDF5 group containing the instrument data and metadata for the SWIR3 strip. The structure and content of this group is elaborated in Table 23.	See Table 22.
HDF5 Attribute	ICP_REFERENCE	Reference to the used radiometric ICP file. This string has following syntax: PROBAV_ICP_RADIOMETRIC#{LEFT|RIGHT|CENTER}_<StartDate>_V<Revision>	String
HDF5 Attribute	PROCESSINGINFO_RADIOMODELLING	Reference to the used radio modelling tool.	String

Table 23: HDF5 structure of LEVEL 1C STRIP (BLUE, NIR, RED, SWIR1, SWIR2, and SWIR3) Groups.

Type	Name	Description	Data type
HDF5 Dataset	Q	Dataset containing the quality indicator values. Every pixel is decoded in an 8-bit integer value. • status = 0 'Correct': no issues • status = 1 'Missing' : the pixel value is missing due to a bad detector • status = 2 'WasSaturated' : the value DN of the pixel is equal to 4095 (coded in 12 bits = 2^12-1) • status = 3 'BecameSaturated' : during the calculation of the TOA reflectance, the value becomes higher than 4095 • status = 4 'BecameNegative' : during the calculation of the TOA reflectance, the value becomes lower than 0 • status = 5 'Interpolated' : the value of the radiance of the pixel is interpolated using the neighbour pixels • status = 6 'BorderCompressed': the quality of theses pixels of a strip are uncertain due to on-board compression artefacts. Table 24 lists the metadata items specific for this dataset.	See Table 24
HDF5 Dataset	TOA	Dataset containing the Top-Of-Atmosphere (TOA) reflectance values. Table 24 lists the metadata items specific for this dataset.	16-bit integer
HDF5 Attribute	DETECTOR	Identifier for the detector. Possible values are: VNIR, SWIR	String

Type	Name	Description	Data type
HDF5 Attribute	OBSERVATION_END_DATE	The observation end date (UTC), i.e. the date of the last line of the strip. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_END_TIME	The observation end time (UTC), i.e. the time of the last line of the strip. The format is: hh:mm:ss.µµµµµµ.	String
HDF5 Attribute	OBSERVATION_START_DATE	The observation start date (UTC), i.e. the date of the first line of the strip. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_START_TIME	The observation start time (UTC), i.e. the time of the first line of the strip. The format is: hh:mm:ss.µµµµµµ.	String
HDF5 Attribute	SOLAR_IRRADIANCE	The solar Top-Of-Atmosphere irradiance [W m^{-2}].	32-bit floating point
HDF5 Attribute	STRIP	Identifier for the strip. Possible values are: BUE, RED, NIR, SWIR1 (left SWIR strip), SWIR2 (center SWIR strip), SWIR3 (right SWIR strip)	String

Table 24: HDF5 metadata items for the LEVEL 1C attributes.

Type	Name	Description	Data type
HDF5 Attribute	DESCRIPTION	Short description of the dataset.	String
HDF5 Attribute	MAPPING	The mapping information, consisting of following values: • **<proj_id>**: the projection ID (e.g. “Geographic Lat/Lon”) • **<x_m>**: A value indicating whether the map X coordinates refer to the top-left corner (0.0) or center (0.5) of the pixel. • **<y_m>**: A value indicating whether the map Y coordinates refer to the top-left corner (0.0) or center (0.5) of the pixel. • **<x_start>**: the X coordinate of the upper-left pixel. • **<y_start>**: the Y coordinate of the upper-left pixel. • **<x_res>**: the spatial resolution in the X direction. • **<y_res>**: the spatial resolution in the Y direction. • **<datum>**: the projection‘s datum (in case of unprojected image, the value is ‘-’). • **<units>**: the projection ‘s unit (in case of unprojected image, the value is ‘-’). Note that his is an optional attribute. If not provided, it is assumed that no geographical information is attached to the dataset.	String
HDF5 Attribute	NO_DATA	The no data value.	32-bit floating-point
HDF5 Attribute	OFFSET	The offset factor. The physical value (PV) is calculated as PV = (DN-offset)/scale (DN = digital number)	32-bit floating-point
HDF5 Attribute	SCALE	The scale factor. The physical value (PV) is calculated as PV = (DN-offset)/scale (DN = digital number)	32-bit floating-point
HDF5 Attribute	UNITS	The units.	String

Appendix E2: Detailed Level 2A Product file description

Below follows a detailed description of the various Groups, Datasets, and Attributes of Level 2A files is given. Reference is made to Figure 19, in which the HDF5 dataset structure is visualised. Note that for the entire Collection 1 CF compliant metadata were added. More information, as well as an overview of the metadata fields is given in Section 4.5 and Table 8, respectively.

Table 25: HDF5 structure of Level 2A file.

Type	Name	Description	Data type
HDF5 Group	LEVEL 2A	HDF5 "root" group containing the Level 2A TOA/TOC data and metadata. The structure and content of this group is elaborated in Table 33.	-
HDF5 Attribute	ACQUISITION_ STATION	Name of the acquisition station, default value is 'Kiruna'.	String
HDF5 Attribute	CAMERA	Camera identifier. Possible values are: "LEFT" for the left camera (camera 1), "CENTER" for the center camera (camera 2), "RIGHT" for the right camera (camera 3)	String
HDF5 Attribute	DATELINE_ CROSSING	Flag indicating whether or not the segment is crossing the International Dateline.	String
HDF5 Attribute	DAY_NIGHT_FLAG	Indicating whether or not the segment is a day segment or taken at night.	String
HDF5 Attribute	DESCRIPTION	Short description of the file content, e.g. PROBA-V Level2A S1 Top-Of-Atmosphere product at 1 km	String
HDF5 Attribute	INSTRUMENT	Short name for the instrument, i.e., VEGETATION.	String
HDF5 Attribute	MAP_PROJECTION_ FAMILY	The family to which the projection belongs.	String
HDF5 Attribute	MAP_PROJECTION_ NAME	The full name of the projection.	String
HDF5 Attribute	MAP_PROJECTION_ REFERENCE	A unique reference (EPSG code) of the projection. An example is EPSG:4326.	String
HDF5 Attribute	MAP_PROJECTION_ UNITS	The units of the projection. Possible values are: DEGREES, METERS	String
HDF5 Attribute	MAP_PROJECTION_ WKT	The projection WKT string.	String
HDF5 Attribute	NORTHPOLE_ INDICATOR	Flag indicating whether or not the segment is covering the north pole.	String
HDF5 Attribute	OBSERVATION_ END_DATE	The observation end date (UTC) of the segment. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_ END_TIME	The observation end time (UTC) of the segment. The format is: hh:mm:ss.μμμμμμ.	String
HDF5 Attribute	OBSERVATION_ START_DATE	The observation start date (UTC) of the segment. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_ START_TIME	The observation start time (UTC) of the segment. The format is: hh:mm:ss.μμμμμμ.	String
HDF5 Attribute	OVERPASS_	The overpass number since launch.	32-bit integer

Type	Name	Description	Data type
	NUMBER		
HDF5 Attribute	PLATFORM	Short name for the platform and its serial number, i.e. PROBA-1.	String
HDF5 Attribute	PROCESSING_DATE	The date the product was generated. The format is: YYYY-MM-DD.	String
HDF5 Attribute	PROCESSING_TIME	The time the product was generated. The format is: hh:mm:ss.μμμμμμ.	String
HDF5 Attribute	PRODUCT_ REFERENCE	A unique textual reference to the product. This string has following syntax: Segment_PROBAV#<YYYYMMDDhhmmss>_<LEVEL>_<RESOLUTION>_V<VERSION> Where: <YYYYMMDDhhmmss> is the start observation date and time; <LEVEL> is L2A, <RESOLUTION> is the spatial resolution of the data, and <VERSION> is the version identifier (three digits).	String
HDF5 Attribute	VERSION	Denotes the product version	32-bit integer
HDF5 Attribute	SYNTHESIS_PERIOD	The synthesis period. Following values are possible: 1: for a daily synthesis, 10: for a 10-day synthesis	32-bit integer
HDF5 Attribute	SOUTHPOLE_INDICATOR	Flag indicating whether or not the segment is covering the south pole.	String

Table 26: HDF5 structure of LEVEL 2A Root Group.

Type	Name	Description	Data type
HDF5 Group	GEOMETRY	HDF5 group containing the geometry data for the segment. The structure and content of this group is elaborated in Table 34.	
HDF5 Group	QUALITY	HDF5 group containing the quality data for the segment. The structure and content of this group is elaborated in Table 36.	
HDF5 Group	RADIOMETRY	HDF5 group containing the radiometry data for the segment. The structure and content of this group is elaborated in Table 37.	
HDF5 Attribute	GEOMETRIC_ICP_REFERENCE	Reference to the used geometric ICP file. This string has the following syntax: PROBAV_ICP_GEOMETRIC#{LEFT|RIGHT|CENTER}_<StartDate>_V<Revision>	String
HDF5 Attribute	PROCESSINGINFO_ CLOUDICESNOW_ DETECTION	Reference to the used cloud, snow and ice detection algorithm version.	String
HDF5 Attribute	PROCESSINGINFO_ GEOMODELLING	Reference to the used geo modelling algorithm version.	String
HDF5 Attribute	PROCESSINGINFO_ MAPPING	Reference to the used mapping algorithm version.	String
HDF5 Attribute	PROCESSINGINFO_ MOSAIC	Reference to the used mosaicking algorithm version.	String

Type	Name	Description	Data type
HDF5 Attribute	PROCESSINGINFO_ RADIOMODELLING	Reference to the used radio modelling algorithm version.	String
HDF5 Attribute	PROCESSINGINFO_ SHADOWDETECTION	Reference to the used shadow detection algorithm version.	String
HDF5 Attribute	RADIOMETRIC_ICP_REFERENCE	Reference to the used radiometric ICP file. This string has following syntax: PROBAV_ICP_RADIOMETRIC#{LEFT|RIGHT|CENTER}_<StartDate>_V<Revision>	String

Table 27: HDF5 structure of GEOMETRY group.

Type	Name	Description	Data type
HDF5 Group	SWIR	HDF5 group containing the data and metadata for the SWIR detector. This group contains two datasets: • **VAA**: dataset containing the viewing azimuth angles. Every pixel is decoded in an 8-bit unsigned integer value. • **VZA**: dataset containing the viewing zenith angles. Every pixel is decoded in an 8-bit unsigned integer value. Table 40 lists the metadata items specific for this dataset.	-
HDF5 Group	VNIR	HDF5 group containing the data and metadata for the VNIR detector. This group contains two datasets: • **VAA**: dataset containing the viewing azimuth angles. Every pixel is decoded in an 8-bit unsigned integer value. • **VZA**: dataset containing the viewing zenith angles. Every pixel is decoded in an 8-bit unsigned integer value. Table 40 lists the metadata items specific for this dataset	-
HDF5 Dataset	CONTOUR	Compound dataset containing latitude and longitude information.	
HDF5 Dataset	SAA	Dataset containing the solar azimuth angles. Every pixel is decoded as an 8-bit unsigned integer value. Table 40 lists the metadata items specific for this dataset.	-
HDF5 Dataset	SZA	HDF5 dataset containing the SZA. Every pixel is decoded as an 8-bit unsigned integer value. Table 40 lists the metadata items specific for this dataset.	-
HDF5 Attribute	BOTTOM_LEFT_LATITUDE	The latitude of the bottom-left point of the segment.	32-bit floating-point

Type	Name	Description	Data type
HDF5 Attribute	BOTTOM_LEFT_LONGITUDE	The longitude of the bottom-left point of the segment.	32-bit floating-point
HDF5 Attribute	BOTTOM_LEFT_X	The X coordinate of the bottom-left point of the cartographic bounding box of the segment.	32-bit floating-point
HDF5 Attribute	BOTTOM_LEFT_Y	The Y coordinate of the bottom-left point of the cartographic bounding box of the segment.	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_LATITUDE	The latitude of the bottom-right point of the segment.	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_LONGITUDE	The longitude of the bottom-right point of the segment.	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_X	The X coordinate of the bottom-right point of the cartographic bounding box of the segment.	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_Y	The Y coordinate of the bottom-right point of the cartographic bounding box of the segment.	32-bit floating-point
HDF5 Attribute	CENTER_LATITUDE	The latitude of the center point of the geographic bounding box of the segment.	32-bit floating-point
HDF5 Attribute	CENTER_LONGITUDE	The longitude of the center point of the geographic bounding box of the segment.	32-bit floating-point
HDF5 Attribute	CENTER_X	The center point in X direction of the cartographic bounding box of the segment.	32-bit floating-point
HDF5 Attribute	CENTER_Y	The center point in Y direction of the cartographic bounding box of the segment.	32-bit floating-point
HDF5 Attribute	TOP_LEFT_LATITUDE	The latitude of the top-left point of the segment.	32-bit floating-point
HDF5 Attribute	TOP_LEFT_LONGITUDE	The longitude of the top-left point of the segment.	32-bit floating-point
HDF5 Attribute	TOP_LEFT_X	The X - coordinate of the top-left point of the cartographic bounding box of the segment.	32-bit floating-point
HDF5 Attribute	TOP_LEFT_Y	The Y - coordinate of the top-left point of the cartographic bounding box of the segment.	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_LATITUDE	The latitude of the top-right point of the segment.	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_LONGITUDE	The longitude of the top-right point of the segment.	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_X	The X - coordinate of the top-right point of the cartographic bounding box of the segment.	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_Y	The Y - coordinate of the top-right point of the cartographic bounding box of the segment.	32-bit floating-point

Table 28: HDF5 structure of QUALITY Group.

<table>
<tr><th>Type</th><th>Name</th><th>Description</th><th>Data type</th></tr>
<tr><td>HDF5 Dataset</td><td>SM</td><td>Dataset containing the quality flags and status pixels.
<table>
<tr><th>Bit (LSB to MSB)</th><th>Description</th><th>Value</th><th>Key</th></tr>
<tr><td>0-2</td><td>Cloud/Ice Snow/Shadow Flag</td><td>000
001
010
011
100</td><td>Clear
Shadow
Undefined
Cloud
Ice</td></tr>
<tr><td>3</td><td>Land/Sea</td><td>0
1</td><td>Sea
Land</td></tr>
<tr><td>4</td><td>Radiometric quality SWIR flag</td><td>0
1</td><td>Bad
Good</td></tr>
<tr><td>5</td><td>Radiometric quality NIR flag</td><td>0
1</td><td>Bad
Good</td></tr>
<tr><td>6</td><td>Radiometric quality RED flag</td><td>0
1</td><td>Bad
Good</td></tr>
<tr><td>7</td><td>Radiometric quality BLUE flag</td><td>0
1</td><td>Bad
Good</td></tr>
<tr><td>8</td><td>SWIR coverage</td><td>0
1</td><td>No
Yes</td></tr>
<tr><td>9</td><td>NIR coverage</td><td>0
1</td><td>No
Yes</td></tr>
<tr><td>10</td><td>RED coverage</td><td>0
1</td><td>No
Yes</td></tr>
<tr><td>11</td><td>BLUE coverage</td><td>0
1</td><td>No
Yes</td></tr>
</table>
Table 36 lists the metadata items specific for this dataset.</td><td>8-bit unsigned integer</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>HDF5 Attribute</td><td>PERCENTAGE _CLOUD</td><td>The percentage cloud.</td><td>32-bit floating-point</td></tr>
<tr><td>HDF5 Attribute</td><td>PERCENTAGE _SNOW</td><td>The percentage snow.</td><td>32-bit floating-point</td></tr>
<tr><td>HDF5 Attribute</td><td>PERCENTAGE _LAND</td><td>The percentage land.</td><td>32-bit floating-point</td></tr>
<tr><td>HDF5 Attribute</td><td>PERCENTAGE _MISSING_ DATA</td><td>The percentage missing data.</td><td>32-bit floating-point</td></tr>
</table>

Table 29: HDF5 structure of RADIOMETRY Group.

Type	Name	Description
HDF5 Group	BLUE	HDF5 group containing the radiometry data for BLUE band of the segment. The structure and content of this group is explained in Table 30.
HDF5 Group	NIR	HDF5 group containing the radiometry data for NIR band of the segment. The structure and content of this group is explained in Table 30.
HDF5 Group	RED	HDF5 group containing the radiometry data for RED band of the segment. The structure and content of this group is explained in Table 30
HDF5 Group	SWIR	HDF5 group containing the radiometry data for SWIR band of the segment. The structure and content of this group is explained in Table 30.

Table 30: HDF5 structure of band groups in the RADIOMETRY Group.

Type	Name	Description	Data type
HDF5 Dataset	TOA	Dataset containing the TOA reflectances. Table 31 lists the metadata items specific for this dataset.	16-bit integer
HDF5 Attribute	DETECTOR	Identifier for the detector. Possible values are: VNIR, SWIR	String
HDF5 Attribute	GAIN_FACTOR	The gain factor.	32-bit floating-point
HDF5 Attribute	OBSERVATION_END_DATE	The observation end date (UTC), i.e. the date of the last line of the band. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_END_TIME	The observation end time (UTC), i.e. the time of the last line of the band. The format is: hh:mm:ss.µµµµµµ .	String
HDF5 Attribute	OBSERVATION_START_DATE	The observation start date (UTC), i.e. the date of the first line of the band. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_START_TIME	The observation start time (UTC), i.e. the time of the first line of the band. The format is: hh:mm:ss.µµµµµµ.	String
HDF5 Attribute	SOLAR_IRRADIANCE	The solar irradiance at TOA for the respective band.	32-bit floating-point

Table 31: HDF5 metadata items for the datasets.

Type	Name	Description	Data type
HDF5 Attribute	DESCRIPTION	Short description of the dataset.	String
HDF5 Attribute	DIMENSION_ LABEL	Lat, lon	String
HDF5 Attribute	DIMENSION_ LIST	Arrays with object references to other datasets.	Variable-length of Object reference
HDF5 Attribute	MAPPING	The mapping information, consisting of following values: • **<proj_id>**: the projection ID (e.g. “Geographic Lat/Lon”) • **<x_m>**: A value indicating whether the map X coordinates refer to the top-left corner (0.0) or center (0.5) of the pixel. • **<y_m>**: A value indicating whether the map Y coordinates refer to the top-left corner (0.0) or center (0.5) of the pixel. • **<x_start>**: the X coordinate of the upper-left pixel. • **<y_start>**: the Y coordinate of the upper-left pixel. • **<x_res>**: the spatial resolution in the X direction. • **<y_res>**: the spatial resolution in the Y direction. • **<datum>**: the projection‘s datum (in case of unprojected image, the value is ‘-’). • **<units>**: the projection ‘s unit (in case of unprojected image, the value is ‘-’). Note that his is an optional attribute. If not provided, it is assumed that no geographical information is attached to the dataset.	String
HDF5 Attribute	NO_DATA	The no data value.	32-bit floating-point
HDF5 Attribute	OFFSET	The scale factor. The physical value (PV) is calculated as PV = (DN-offset)/scale (DN = Digital Number Count)	32-bit floating-point
HDF5 Attribute	SCALE	The scale factor. The physical value (PV) is calculated as PV = (DN-offset)/scale (DN = Digital Number Count)	32-bit floating-point
HDF5 Attribute	UNITS	The units of the dataset.	String

Appendix E3: Detailed Synthesis (S1/S5/S10) product file description

Below follows the detailed description of the various Groups, Datasets, and Attributes of the Synthesis product files. Reference is made to Figure 21, in which the HDF5 dataset structure is visualized. Note that for the entire Collection 1 CF compliant metadata were added. More information, as well as an overview of the metadata fields is given in Section 4.5 and Table 8, respectively.

Table 32: HDF5 structure of Synthesis file.

Type	Name	Description	Data type
HDF5 Group	LEVEL3	HDF5 “root” group containing the Level3 TOA/TOC data and metadata. The structure and content of this group is elaborated in Table 33.	-
HDF5 Attribute	DESCRIPTION	Short description of the file content, e.g. PROBA-V Level3 S1 Top-Of-Atmosphere product at 1km	-
HDF5 Attribute	INSTRUMENT	Short name for the instrument, i.e., VEGETATION.	-
HDF5 Attribute	MAP_PROJECTION_FAMILY	The family to which the projection belongs.	String
HDF5 Attribute	MAP_PROJECTION_NAME	The full name of the projection.	String
HDF5 Attribute	MAP_PROJECTION_REFERENCE	A unique reference (EPSG code) of the projection. An example is EPSG:4326.	String
HDF5 Attribute	MAP_PROJECTION_UNITS	The units of the projection. Possible values are: DEGREES, METERS	String
HDF5 Attribute	MAP_PROJECTION_WKT	The projection string.	String
HDF5 Attribute	PLATFORM	Short name for the platform and its serial number, i.e. PROBA-1.	String
HDF5 Attribute	PROCESSING_DATE	The date the product was generated. The format is: YYYY-MM-DD.	String
HDF5 Attribute	PROCESSING_TIME	The time the product was generated. The format is: hh:mm:ss.µµµµµµ.	String
HDF5 Attribute	PRODUCT_REFERENCE	A unique textual reference to the product. This string has following syntax: Synthesis_PROBAV_<YYYYMMDD>_<LEVEL> _ <GRID>_V<VERSION> Where: <YYYYMMDD> is the start observation date; <LEVEL> is ‘S1_TOA’, ‘S1_TOC’ or ‘S10_TOC’; <GRID> is the spatial resolution; <VERSION> is the version identifier (three digits)	String
HDF5 Attribute	VERSION	Denotes the product version	32-bit integer
HDF5 Attribute	SYNTHESIS_PERIOD	The synthesis period. Following values are possible: 1: daily synthesis, 5: 5-day synthesis, 10: 10-daily synthesis	32-bit integer

Table 33: HDF5 structure of LEVEL3 Root Group.

Type	Name	Description	Data type
HDF5 Group	GEOMETRY	HDF5 group containing the geometry data for the synthesis. The structure and content of this group is elaborated in Table 34.	-
HDF5 Group	NDVI	HDF5 group containing the NDVI (Normalized Difference Vegetation Index) data for the synthesis. The structure and content of this group is elaborated in Table 35.	-
HDF5 Group	QUALITY	HDF5 group containing the quality data for the synthesis. The structure and content of this group is elaborated in Table 36.	-
HDF5 Group	RADIOMETRY	HDF5 group containing the radiometry data for the synthesis. The structure and content of this group is elaborated in Table 37.	-
HDF5 Group	TIME	HDF5 group containing the timing data for the synthesis. The structure and content of this group is elaborated in Table 39.	-
HDF5 Attribute	PROCESSINGINFO_ CLOUDICESNOW_ DETECTION	Reference to the used cloud, snow and ice detection algorithm version.	String
HDF5 Attribute	PROCESSINGINFO_ COMPOSITING	Reference to the used compositing algorithm version.	String
HDF5 Attribute	PROCESSINGINFO_ GEOMODELLING	Reference to the used geo modelling algorithm version.	String
HDF5 Attribute	PROCESSINGINFO_ MAPPING	Reference to the used mapping algorithm version.	String
HDF5 Attribute	PROCESSINGINFO_ MOSAIC	Reference to the used mosaicking algorithm version.	String
HDF5 Attribute	PROCESSINGINFO_ RADIOMODELLING	Reference to the used radio modelling algorithm version.	String
HDF5 Attribute	PROCESSINGINFO_ SHADOWDETECTION	Reference to the used shadow detection algorithm version.	String

Table 34: HDF5 structure of GEOMETRY group.

Type	Name	Description	Data type
HDF5 Group	SWIR	HDF5 group containing the data and metadata for the SWIR – detector. This group contains two datasets: • **VAA**: dataset containing the viewing azimuth angles. Every pixel is decoded in an 8-bit unsigned integer value. • **VZA**: dataset containing the viewing zenith angles. Every pixel is decoded in an 8-bit unsigned integer value. Table 40 lists the metadata items specific for this dataset.	
HDF5 Group	VNIReserved	HDF5 group containing the data and metadata for the VNIR detector. This group contains two datasets: • **VAA**: dataset containing the viewing azimuth angles. Every pixel is decoded in an 8-bit unsigned integer value. • **VZA**: dataset containing the viewing zenith angles. Every pixel is decoded in an 8-bit unsigned integer value. Table 40 lists the metadata items specific for this dataset	
HDF5 Dataset	SAA	Dataset containing the solar azimuth angles. Every pixel is decoded as an 8-bit unsigned integer value. Table 40 lists the metadata items specific for this dataset.	
HDF5 Dataset	SZA	HDF5 dataset containing the SZA. Every pixel is decoded as an 8-bit unsigned integer value. Table 40 lists the metadata items specific for this dataset.	
HDF5 Attribute	BOTTOM_LEFT_LATITUDE	The latitude of the bottom-left point of the synthesis.	32-bit floating-point
HDF5 Attribute	BOTTOM_LEFT_LONGITUDE	The longitude of the bottom-left point of the synthesis.	32-bit floating-point
HDF5 Attribute	BOTTOM_LEFT_X	The X coordinate of the bottom-left point of the cartographic bounding box of the synthesis.	32-bit floating-point
HDF5 Attribute	BOTTOM_LEFT_Y	The Y coordinate of the bottom-left point of the cartographic bounding box of the synthesis.	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_LATITUDE	The latitude of the bottom-right point of the synthesis.	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_LONGITUDE	The longitude of the bottom-right point of the synthesis.	32-bit floating-point

Type	Name	Description	Data type
HDF5 Attribute	BOTTOM_RIGHT_X	The X coordinate of the bottom-right point of the cartographic bounding box of the synthesis.	32-bit floating-point
HDF5 Attribute	BOTTOM_RIGHT_Y	The Y coordinate of the bottom-right point of the cartographic bounding box of the synthesis.	32-bit floating-point
HDF5 Attribute	CENTER_LATITUDE	The latitude of the center point of the geographic bounding box of the synthesis.	32-bit floating-point
HDF5 Attribute	CENTER_LONGITUDE	The longitude of the center point of the geographic bounding box of the synthesis.	32-bit floating-point
HDF5 Attribute	CENTER_X	The center point in X direction of the cartographic bounding box of the synthesis.	32-bit floating-point
HDF5 Attribute	CENTER_Y	The center point in Y direction of the cartographic bounding box of the synthesis.	32-bit floating-point
HDF5 Attribute	TOP_LEFT_LATITUDE	The latitude of the top-left point of the synthesis.	32-bit floating-point
HDF5 Attribute	TOP_LEFT_LONGITUDE	The longitude of the top-left point of the synthesis.	32-bit floating-point
HDF5 Attribute	TOP_LEFT_X	The X - coordinate of the top-left point of the cartographic bounding box of the synthesis.	32-bit floating-point
HDF5 Attribute	TOP_LEFT_Y	The Y - coordinate of the top-left point of the cartographic bounding box of the synthesis.	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_LATITUDE	The latitude of the top-right point of the synthesis.	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_LONGITUDE	The longitude of the top-right point of the synthesis.	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_X	The X - coordinate of the top-right point of the cartographic bounding box of the synthesis.	32-bit floating-point
HDF5 Attribute	TOP_RIGHT_Y	The Y - coordinate of the top-right point of the cartographic bounding box of the synthesis.	32-bit floating-point

Table 35: HDF5 structure of NDVI Group.

Type	Name	Description	Data type
HDF5 Dataset	NDVI	Dataset containing the NDVI (Normalized Difference Vegetation Index). Table 40 lists the metadata items specific for this dataset.	8-bit unsigned integer

Table 36: HDF5 structure of QUALITY Group.

<table>
<tr><th>Type</th><th>Name</th><th>Description</th><th>Data type</th></tr>
<tr><td>HDF5 Dataset</td><td>SM</td><td>Dataset containing the quality flags and status pixels.
<table>
<tr><th>Bit (LSB to MSB)</th><th>Description</th><th>Value</th><th>Key</th></tr>
<tr><td>0-2</td><td>Cloud/Ice Snow/Shadow Flag</td><td>000
001
010
011
100</td><td>Clear
Shadow
Undefined
Cloud
Ice</td></tr>
<tr><td>3</td><td>Land/Sea</td><td>0
1</td><td>Sea
Land</td></tr>
<tr><td>4</td><td>Radiometric quality SWIR flag</td><td>0
1</td><td>Bad
Good</td></tr>
<tr><td>5</td><td>Radiometric quality NIR flag</td><td>0
1</td><td>Bad
Good</td></tr>
<tr><td>6</td><td>Radiometric quality RED flag</td><td>0
1</td><td>Bad
Good</td></tr>
<tr><td>7</td><td>Radiometric quality BLUE flag</td><td>0
1</td><td>Bad
Good</td></tr>
</table>
Table 40 lists the metadata items specific for this dataset.</td><td>8-bit unsigned integer</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>HDF5 Attribute</td><td>PERCENTAGE_ CLOUD</td><td>The percentage cloud.</td><td>32-bit floating-point</td></tr>
<tr><td>HDF5 Attribute</td><td>PERCENTAGE_ SNOW</td><td>The percentage snow.</td><td>32-bit floating-point</td></tr>
<tr><td>HDF5 Attribute</td><td>PERCENTAGE_ LAND</td><td>The percentage land.</td><td>32-bit floating-point</td></tr>
<tr><td>HDF5 Attribute</td><td>PERCENTAGE_ MISSING_ DATA</td><td>The percentage missing data.</td><td>32-bit floating-point</td></tr>
</table>

Table 37: HDF5 structure of RADIOMETRY Group.

Type	Name	Description
HDF5 Group	BLUE	HDF5 group containing the radiometry data for BLUE band of the synthesis. The structure and content of this group is explained in Table 38.
HDF5 Group	NIR	HDF5 group containing the radiometry data for NIR band of the synthesis. The structure and content of this group is explained in Table 38.
HDF5 Group	RED	HDF5 group containing the radiometry data for RED band of the synthesis. The structure and content of this group is explained in Table 38.
HDF5 Group	SWIR	HDF5 group containing the radiometry data for SWIR band of the synthesis. The structure and content of this group is explained in Table 38.

Table 38: HDF5 structure of band groups in the RADIOMETRY Group.

Type	Name	Description	Data type
HDF5 Dataset	TOA or TOC	Dataset containing the Top-Of-Atmosphere reflectance values (TOA) or Top-Of-Canopy reflectance values (TOC). Table 40 lists the metadata items specific for this dataset.	16-bit integer
HDF5 Attribute	DETECTOR	Identifier for the detector (type: string). Possible values are: VNIR, SWIR	String
HDF5 Attribute	GAIN_FACTOR	The gain factor (type: float).	32-bit floating-point
HDF5 Attribute	OBSERVATION_END_DATE	The observation end date (UTC), i.e. the date of the last line of the band (type: string). The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_END_TIME	The observation end time (UTC), i.e. the time of the last line of the band (type: string). The format is: hh:mm:ss.µµµµµµ .	String
HDF5 Attribute	OBSERVATION_START_DATE	The observation start date (UTC), i.e. the date of the first line of the band (type: string). The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_START_TIME	The observation start time (UTC), i.e. the time of the first line of the band (type: string). The format is: hh:mm:ss.µµµµµµ.	String
HDF5 Attribute	SOLAR_IRRADIANCE	The solar irradiance at TOA.	32-bit floating-point

Table 39: HDF5 structure of TIME Group.

Type	Name	Description	Data type
HDF5 Dataset	TIME	Dataset containing the start acquisition time of the selected segment, expressed in minutes since the beginning of the synthesis period in UTC. Table 40 lists the metadata items specific for this dataset.	16-bit unsigned integer
HDF5 Attribute	OBSERVATION_END_DATE	The observation end date (UTC) of the synthesis. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_END_TIME	The observation end time (UTC) of the synthesis. The format is: hh:mm:ss.	String
HDF5 Attribute	OBSERVATION_START_DATE	The observation start date (UTC) of the synthesis. The format is: YYYY-MM-DD.	String
HDF5 Attribute	OBSERVATION_START_TIME	The observation start time (UTC) of the synthesis. The format is: hh:mm:ss.	String

Table 40: HDF5 metadata items for the datasets.

Type	Name	Description	Data type
HDF5 Attribute	DESCRIPTION	Short description of the dataset.	String
HDF5 Attribute	DIMENSION_ LABEL	Lat, lon	String
HDF5 Attribute	DIMENSION_ LIST	Arrays with object references to other datasets.	Variable-length of Object reference
HDF5 Attribute	MAPPING	The mapping information, consisting of following values: • **<proj_id>**: the projection ID (e.g. "Geographic Lat/Lon") • **<x_m>**: A value indicating whether the map X coordinates refer to the top-left corner (0.0) or center (0.5) of the pixel. • **<y_m>**: A value indicating whether the map Y coordinates refer to the top-left corner (0.0) or center (0.5) of the pixel. • **<x_start>**: the X coordinate of the upper-left pixel. • **<y_start>**: the Y coordinate of the upper-left pixel. • **<x_res>**: the spatial resolution in the X direction. • **<y_res>**: the spatial resolution in the Y direction. • **<datum>**: the projection's datum (in case of unprojected image, the value is '-'). • **<units>**: the projection 's unit (in case of unprojected image, the value is '-'). Note that his is an optional attribute. If not provided, it is assumed that no geographical information is attached to the dataset.	String
HDF5 Attribute	NO_DATA	The no data value.	64-bit floating-point
HDF5 Attribute	OFFSET	The scale factor. The physical value (PV) is calculated as PV = (DN-offset)/scale (DN = Digital Number Count)	32-bit floating-point
HDF5 Attribute	SCALE	The scale factor. The physical value (PV) is calculated as PV = (DN-offset)/scale (DN = Digital Number Count)	32-bit floating-point
HDF5 Attribute	UNITS	The units of the dataset.	String