S2 MPC

Sen2like User Manual

Ref. S2-SEN2LIKE-UM-V1.4

Authors Table

	Name	Company	Responsibility	Date	Signature
Written by	S. Saunier	TPZF	Sen2Like Team	2020-10-14	
Verified by	O Devignot	CS	Quality Manager	2020-10-14	
Approved by	L. Pessiot	CS	Service Manager	2020-10-14	

Change Log

Issue	Date	Reason for change	Pages(s)/Section(s)
1 Draft A	06 March 2019	Initial Issue.(HLS TPZF S2L UM0017)	All
	26 May 2020	Major update in the CCN context	All
1.1	23 June 2020	Final version	All
1.2	14 Oct. 2020	Update with new features	All
1.3	21 Sept. 2021	Update for release v3.2.0 - Maja product Format specification - To clarify band Parallelisation - Reference LS09	All
1.4	07 Dec. 2021	Update for release dedicated to massive production	All

This Page Is Intentionally Blank

Table of contents

This Page Is Intentionally Blank

1. Introduction

1.1 Purpose of the document

This document is the user manual document of the "Sen2Like" phase 2 software, developed in the frame of the two following S2 Mission Performance Centre (MPC) Contract Change Requests; Sen2Like phase 1 [ESA-EOPG-Cop-CR-1], Sen2Like phase 2 [ESA-EOPG-Cop-CR-11]; contract ref. 4000108650/13/I-LG.

1.2 Document structure

The document is structured as follows:

- Chapter 1 – This introduction
- Chapter 2 – This chapter describes the tool.
- Chapter 3 – This chapter provides the installation note of the software package.
- Chapter 4 – This chapter details the command lines for proper execution.
- Chapter 5 – The chapter address all aspects related to the configuration: Processor & Auxiliary data.

1.3 References

The reference list of all project related documents with their version number and issue date is given in:

[RD.1] Sen2Like, a tool to generate Sentinel-2 Harmonised Surface Reflectance Products, First Results With Landsat-8, 3rd S2 Validation Team Meeting[1]

[RD.2] Harmonized Landsat-8 Sentinel-2 (HLS) Product User's Guide, December 2018 https://hls.gsfc.nasa.gov/wp-content/uploads/2019/01/HLS.v1.4.UserGuide_draft_ver3.1.pdf

[RD.3] S. Saunier et al., "Sen2Like, A Tool To Generate Sentinel-2 Harmonised Surface Reflectance Products - First Results with Landsat-8," IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019, pp. 5650-5653, doi: 10.1109/IGARSS.2019.8899213.

[RD.4] Landsat 8 Level 1 Data Format Control Book (DFCB), Version 11.0, February 2017[2]

---

[1] https://www.researchgate.net/publication/332428332_Sen2like_a_Tool_to_Generate_Sentinel-2_Harmonised_Surface_Reflectance_Products_-_First_Results_With_Landsat-8

[2] https://prd-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/atoms/files/LSDS-809-Landsat8-Level1DFCB-v11.pdf

[RD.5] Landsat 8-9 OLI/TIRS Collection 2 Level 2 Data Format Control Book (DFCB), Version 6.0, September 2020

[RD.6] Sentinel-2 Products Specification Document, Version 14.5, March 2018 [S2-PDGS-TAS-DI-PSD]3

[RD.7] Technical Note, Sentinel-2A L2A Product description, Version 1.0, May 2016[4].

[RD.8] Level 2HF Product Format Specification, Version 1.0, December 2020 [S2-PDGS-MPC-L2HF-PFS].

[RD.9] EOGRID Cloud Tool box https://eogrid.esrin.esa.int/cloudtoolbox/

[RD.10] Sen2Cor Atmospheric corrections tool

[RD.11] https://hls.gsfc.nasa.gov/data/v1.4/ HLS Data

[RD.12] Landsat 8 Quality Reports - https://earth.esa.int/web/sppa/mission-performance/esa-3rd-party-missions/landsat-8/oli-tirs/cyclic-quality-reports

[RD.13] Roy and al. (2017). "Examination of Sentinel-2A multi-spectral instrument (MSI) reflectance anisotropy and the suitability of a general method to normalize MSI reflectance to nadir BRDF adjusted reflectance." Remote Sensing of Environment 199 (2017) 25–38

[RD.14] Vermote, E., C.O. Justice, et F.-M. Breon. « Towards a Generalized Approach for Correction of the BRDF Effect in MODIS Directional Reflectances ». IEEE Transactions on Geoscience and Remote Sensing 47, no 3 (mars 2009): 898 908. https://doi.org/10.1109/TGRS.2008.2005977

[RD.15] Franch, B., Vermote, E., Skakun, S., Roger, J. C., Masek, J., Ju, J., Villaescusa-Nadal, J.L. & Santamaria-Artigas, A. (2019). A method for Landsat and Sentinel 2 (HLS) BRDF normalization. Remote Sensing, 11(6), 632.

[RD.16] Claverie, Martin, Junchang Ju, Jeffrey G. Masek, Jennifer L. Dungan, Eric F. Vermote, Jean-Claude Roger, Sergii V. Skakun, et Christopher Justice. « The Harmonized Landsat and Sentinel-2 Surface Reflectance Data Set ». Remote Sensing of Environment 219 (15 décembre 2018): 145-61. https://doi.org/10.1016/j.rse.2018.09.002.

[RD.17] Gao, F, Masek, J, Schwaller, M, Hall, F (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance, IEEE Transactions on Geoscience and Remote Sensing 44, no 8, 2207-2218

1.4 Informative Reference Documents

[3] https://earth.esa.int/documents/247904/685211/Sentinel-2-Products-Specification-Document

[4] https://theia.cnes.fr/atdistrib/documents/PSC-NT-411-0362-CNES_01_00_SENTINEL-2A_L2A_Products_Description.pdf

[ECSS-E-HB-40A]	Software engineering handbook (11 December 2013), https://ecss.nl/hbstms/ecss-e-hb-40a-software-engineering-handbook-11-december-2013

1.5 Relation to other Documents

There are relation with the following documents:

- [SEN2LIKE-PDD], Sen2Like Output Product Format (PDD)
- [SEN2LIKE-PSD], Sen2Like Output Product Format (PSD)
- [SEN2LIKE-VP], Verification plan
- [SEN2LIKE-SDD], Software Design Document
- [SEN2LIKE-ATBD], Algorithm Theoretical Basis Document
- [SEN2LIKE-UM], Software Installation and User Manual
- [SEN2LIKE-TR], Test Report

1.6 Definitions of Terms and Conventions

The following acronyms and abbreviations are used in this report.

API	Application Programming Interface
BOA	Bottom Of Atmosphere
BRDF	Bidirectional Reflectance Distribution Function
CAMS	Copernicus Atmosphere Monitoring Service
CESBIO	Center for Space Studies of BIOsphere
DEM	Digital Elevation Model
ESA	European Space Agency
EU	European Union
KLT	Kanade-Luca-Tomasi
GIPP	Ground Image Processing Parameter
HLS	Harmonized Landsat Sentinel-2
HR	High Resolution
L1	Level 1
L2F	Level 2 Fuzzed (Level-2F)
L2H	Level 2 Harmonized (Level-2H)
MGRS	Military Grid System
MODIS	Moderate Resolution Imaging Spectroradiometer
MPC	Mission Performance Centre
MS	Multi Spectral
MSI	Multi-Spectral Instrument
NASA	National Aeronautics and Space Administration
NBAR	Nadir BRDF-normalized Reflectance

NIR	Near InfraRed
OLI	Operational Land Imager
Pan	Panchromatic
RD	Reference Document
RGN	Red Green Blue
S2A/S2B	Sentinel 2A / 2B
S2L	Sen2Like
SBAF	Spectral Band Adjustment Factor
SCL	Scene Classification (map)
SMAC	Simplified Model for Atmospheric Corrections
SWIR	Short-Wave InfraRed
UTM	Universal Transverse Mercator
VJB	Vermote Justice Breon
WRS	Worldwide Reference System

2. DESCRIPTION

2.1 Main Overview

The Sentinel-2 and Landsat missions have always been of great importance for Earth Observation agricultural applications (Land User / Land Cover) that requires surface reflectance data from Multi-Spectral (MS) High Resolution (HR) instruments.

The scope of Sen2Like is to harmonize Sentinel-2 / Landsat data in order to increase temporal revisit ([RD.1], [RD.2], [RD.3]). The "Sen2Like" term refers to the notion of "Sentinel-2" considered as a reference mission as a baseline principle to produce multi temporal data stack.

In this context, the Sen2Like development goals and main features are:

- A data harmonization framework,
- Three production modes: 'Product', 'Single Tile' and 'Multi Tile'
- Option to produce in near real time context and in cloud environment (DIAS)
- Delivery of two product types: harmonized ("Level-2H") and fused ("Level-2F") data products

The Level 2F includes Blue, Green, Red Landsat 8 image bands rescaled to 10.0 m pixel spacing.

The Sen2Like software is an open source solution. Designed as an Earth Observation data demonstration processor, the software ingests Sentinel 2 (Level 1, Level 2)/ Landsat 8 / Landsat 9 (Level 1,level 2) products ([RD.4], [RD.5], [RD.6]), including MAJA products [RD.7].

The software produces, under nominal processing baseline, harmonized product with format described in [RD.8].

As shown in figure below, there are different possible input processing levels depending on the mission. Accordingly, the combination of processing, detailed below, is different. The outputs of the processing are spatio-temporal data stack of Sen2Like products, all image expressed into the Sentinel-2 Military Grid Reference System (MGRS) tiling system.

The product format can be either Level-2H or Level-2F ([RD.8]). The both formats are quite similar. The difference is that the Level-2H embeds mission dependant harmonized data and the Level 2F embeds mission independent harmonized data: the resolution of all images from equivalent band is the same by following the Sentinel 2 convention.

Whatever the operational mode, the same workflow is followed. In 'multi tile" mode, the user submits an Area of Interest (AOI) as input.

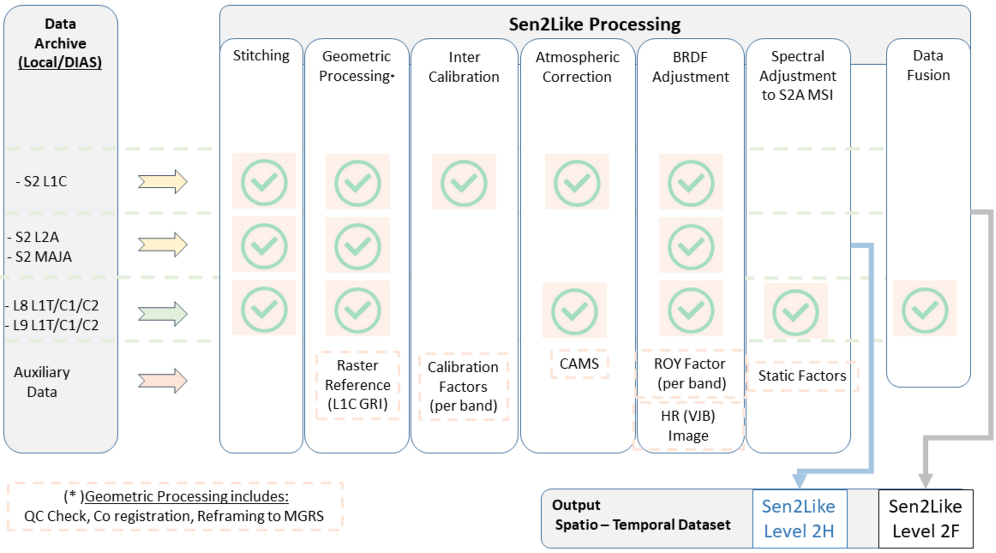

Figure 1 – Overview of Sen2Like processing workflow.

Furthermore, the auxiliary information includes Geometrical Reference data, Digital Elevation Model data (DEM), data from the Copernicus Atmosphere Monitoring Service (CAMS) and geometric reference. Moreover, it optionally includes, Bidirectional Reflectance Distribution Function (BRDF) coefficient data and Sentinel 2 L2A Scene Classification map (SCL).

As discussed above, the Sen2Like band designation convention is defined based on the current Sentinel 2 Multi-Spectral Instrument (MSI) band designation whenever possible. The processing is performed for equivalent Sentinel 2 / Landsat 8 spectral channel images. Even if no match exist, the concerned spectral channel images are kept in the final Sen2Like product and stored in a dedicated folder. It is the case for the Landsat 8 thermal / panchromatic data, and for the Sentinel 2 red edge / NIR1 data.

The table below shows the band naming convention adopted in the L2H / L2F products. In addition, for each spectral band, the resolution of the image is given. The bands for which image fusion algorithm has been applied (L2F) are indicated in bold where corresponding resolution of the image is given. Depending on the band group, the records of the table are displayed with a specific colour. The bands kept as native are indicated in italic. Indeed, mismatch between spectral bands exist; for instance the thermal bands of Landsat 8 do not get equivalency within the Sentinel 2 product. In this case, band file is considered as "native"; related data is stored within a dedicated 'Native' directory, and no sen2lile processing is applied.

The band group nomenclature is listed in

Table 2 and this table is convenient to describe applicability / validity of each processing, as discussed just here after.

Table 1 : Composition of the L2H / L2F products.

Sentinel 2 MSI bands (Center Wavelength [µm])	Lansdat 8 / Landsat 9 bands (Center Wavelength)	Designation	Sen2like Convention	L2H-S2 resolution (m)	L2H-L8 resolution (m)	L2F-S2 resolution (m)	L2F-L8 resolution (m)
B01	B01 (442 nm)	Coastal Aersol	B01	60 m	30 m	60 m	30 m
B02 (490 nm)	B02 (482 nm)	Blue	B02	10 m	**30 m**	10 m	**10 m**
B03 (560nm)	B03 (561 nm)	Green	B03	10 m	**30 m**	10 m	**10 m**
B04 (665 nm)	B04 (654)	Red	B04	10 m	**30 m**	10 m	**10 m**
B08 (842 nm)		*NIR 1*	*B08*	*10 m*	-	*10 m*	-
B8A (865 nm)	B05 (864 nm)	NIR2	B8A	20 m	**30 m**	20 m	**20 m**
B11 (1610 nm)	B06 (1608 nm)	SWIR 1	B11	20 m	**30 m**	20 m	**20 m**
B12 (2190 nm)	B07 (2200 nm)	SWIR 2	B12	20 m	**30 m**	20 m	**20 m**
	B08 (589 nm)	*Panchromatic*	*BP1*	-	*15 m*	-	*15 m*
	B10 (11 µm)	*TIRS 1*	*BT1*	-	*100 m*	-	*100 m*
	B11 (12,2 µm)	*TIRS 2*	*BT2*	-	*100 m*	-	*100 m*
B05 (705 nm)		*Red Edge 1*	*B05*	*20 m*	-	*20 m*	
B06 (740 nm)		*Red Edge 2*	*B06*	*20 m*	-	*20 m*	
B07 (783 nm)		*Red Edge 3*	*B07*	*20 m*	-	*20 m*	

Table 2 : Sen2Like band group convention.

Designation	Band Code Sequence	
Coastal + SWIR	B01, B11,B12	L2F/L2H
RGB	B02,B03,B04	L2F/L2H
NIR	B08,B08A	B8A L2F/L2H B08 “Native”
Pan + Thermal	BP1, BT1, BT2	“Native”
Red Edge	B05,B06,B07	“Native”

For completeness, as given in[5], the definition of the spectral bands are recalled in the graphic below.

[5] http://landsat.gsfc.nasa.gov/wp-content/uploads/2015/06/Landsat.v.Sentinel-2.png

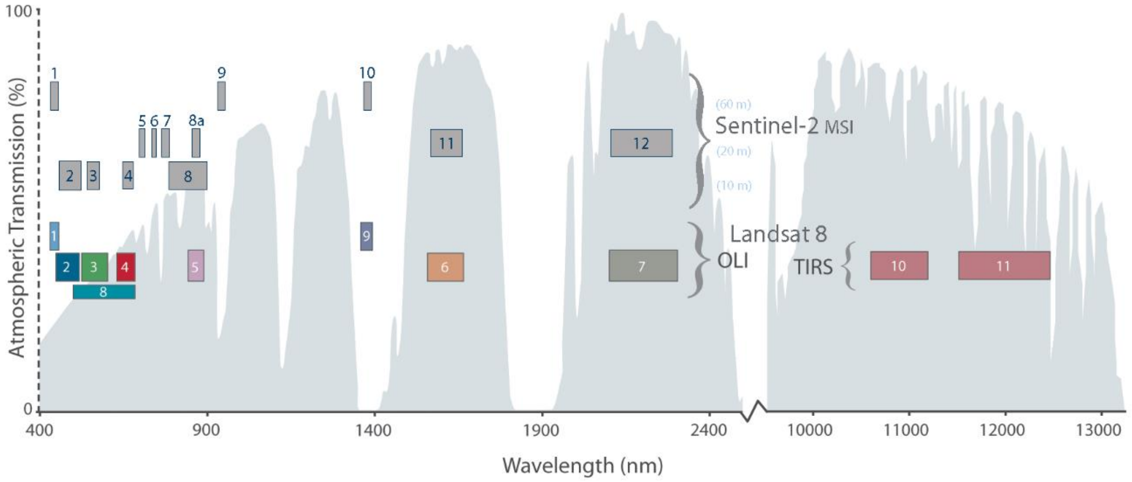

Figure 2 – Sentinel 2 and Landsat 8 spectral band definition.

The Sen2Like prototype processor runs in three following distinct operational modes:

Single tile mode	The processor considers as input one MGRS tile and a period (start time / end time). By using this mode, all corresponding LS8 (LS9) and S2 data are processed. An MGRS multi temporal stack (L2F / L2H) is delivered.
Multi tile mode:	The processor considers as input an Area Of Interest (AOI) and a period of time (start time / end time). By using this mode, all MGRS S2 tiles and LS8 (LS9) scenes that overlap the AOI within a selected time period are processed. An AOI based multi temporal stack (L2F / L2H) is delivered.
Product mode	The processor considers as input only one LS8 (LS9) or S2 product and apply processing for the corresponding MGRS tile.

Regarding the location of input data, there are two ways of running Sen2Like. Classic approach consists in using products stored locally. An alternative approach consists in using products available from the Creodias infrastructure. In this latter case, the catalogue queries are performed with the Creodias opensearch (1.1) like API "Finder"[6]. On the other hand, the data access is done through file system.

[6] https://creodias.eu/eo-data-finder-api-manual

2.2 Processing Algorithms

Sen2Like Phase 2 processor performs the following 7 main processing steps:

- Stitching
- Geometric Processing
- Inter-calibration
- Atmospheric correction
- Bidirectional Reflectance Distribution Function (BRDF) Adjustment
- Spectral Band Adjustment Factor (SBAF)
- Data Fusion

Stitching

In some situations, more than one input scene are required to fully covered the geographical extent of a given MGRS Tile and related images needs to be stitched.

For Landsat, the software selects the two consecutive scenes (the same Worldwide Reference System (WRS)) path and stitched them.

For Sentinel-2, at a given observation date/time, two products are associated to a same MGRS tile when each of them are originating from two distinct data strips[7].

It is assumed that the two image grids are expressed within the same cartographic reference system.

Geometric Processing

In a systematic way, the input images (all spectral bands) are co-registered to a reference image. The reference image needs to be prepared prior to starting the SenLike processing. Note that data from the S2 Global Reference Image (GRI) reference image[8], in the L1C format, can be used as reference, avoiding data preparation stage.

The output projection of Sen2Like products is the Sentinel-2 the tiling system. This tiling system is aligned with the UTM-based MGRS[9].

In geometric processing, both input and reference geometric grids are compared and co-registration errors estimated. An image matching method based on Kanade-Luca-Tomasi (KLT) technics[10], [11], [12] is used. The statistics on geolocation errors provide a correction factors subsequently applied for geometric co registration of input image.

It's worth noting that prior to matching, a dedicated framing process clip the LS8 image to fit within the given MGRS tile geographic extent.

---

[7] https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/definitions

[8] C.Dechoz and al, In proceeding of Living Planet Symposium 2015, Sentinel 2 Global Reference Image

[9] https://en.wikipedia.org/wiki/Military_Grid_Reference_System

[10] Bruce D. Lucas and Takeo Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. International Joint Conference on Artificial Intelligence, pages 674–679, 1981.

[11] Carlo Tomasi and Takeo Kanade. Detection and Tracking of Point Features. Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.

[12] Shi, J.; Tomasi, C. Good Features to Track. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 21–23 June 1994; pp. 593–600

The geometric process ends up with Quality Control; the matching process is played back and the final co-registration accuracy (to reference image) report. The accuracy achieved is within 0.3 pixel (3 sigma).

The co-registration processing does not work on S2 products for which geometric refinement has been applied during the L1C processing (As of Processing Baseline 3.00 release date[13]). This mechanism is overridden by using appropriate parameter as discussed in 5.1.2.5.

Inter-calibration

Radiometric validation activities based on different methods and performed by different teams agree together on mis-calibration between Sentinel-2A (S2A) and Sentinel-2B (S2B): S2B VNIR bands being slightly darker with respect to S2A VNIR bands.

Also, the inter-calibration step aims to bring closer the Top of Atmosphere Radiometry of Sentinel-2B (S2B) with the one of Sentinel-2A (S2A). For this purpose, a scaling factor of 1.011 is applied to the following S2B bands: B01, B02, B03, B04, B05, B06, B07, B08, B8A, B09. Nothing is done for S2B SWIR bands (B10, B11, B12).

With the next generation (scheduled in 2022) of S2 L1C and S2 L2A products (processing baseline > 04.00), it is foreseen to have those inter-calibration coefficient already applied during L1 processing[14].

In Sen2like version 3.3 the inter-calibration coefficients are applied to S2B products if this processing is activated (set to True). It is the responsibility of the user to disable it, in case L1C and L2A products with processing baseline > 04.00 are processed.

The next version of Sen2Like will be able to detect those already inter-calibrated S2B products, in order not to apply the inter-calibration twice.

Atmospheric correction

Starting from Level 1C, it is possible to perform the conversion from Top Of Atmosphere (TOA) to Bottom Of Atmosphere (BOA) by using the two following methods:

- The Simplified Model for Atmospheric Corrections (SMAC) method[15], based on Look up tables (Sensor coefficients are shared by CESBIO)
- The Sen2Cor method[16], as implemented as part of Sentinel 2 Level-2A baseline production.

Both methods relies on auxiliary data provided by European Centre for Medium-Range Weather Forecasts (EMWF), Copernicus Atmosphere Monitoring Service (CAMS) Near Real Time and Reanalysis data[17]. The Simplified Model for

[13] https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/processing-baseline

[14] https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/processing-baseline

[15] Rahman, H., & Dedieu, G. (1994). "SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum." Remote Sens., 15(1), 123-143

[16] Main-Knorn, Magdalena & Pflug, Bringfried & Louis, Jerome & Debaecker, Vincent & Müller-Wilm, Uwe & Gascon, Ferran. (2017). Sen2Cor for Sentinel-2. 3. 10.1117/12.2278218.

[17] https://apps.ecmwf.int/data-catalogues/cams-reanalysis/

Atmospheric Correction (SMAC) method strongly relies on the statistical processing of auxiliary information to estimate Ozone Content, Water Vapour, Temperature, Aerosol Optical Thickness at the time / earth location of the satellite overpass.

Bidirectional Reflectance Distribution Function (BRDF) Adjustment

The viewing and illumination angles are accounted in order to provide nadir BRDF-adjusted reflectance (NBAR). For doing this correction two methods are proposed:

- The c-factors approach: a single set of mission/band dependant coefficients are used to estimate a global BRDF shape as function of viewing, illumination angles, as discussed in given in [RD.13].
- The MGRS Tile BRDF database approach: High resolution BRDF shape characterization is used as detailed in [RD.14] and [RD.15].

The image are normalized per pixel. The view angle is set to nadir and the solar zenith angle is fixed through time but varies for each tile based on latitude as described in [RD.16].It is worth noting that this processing is not available for Cirrus, water vapor, MSI Red Edge and OLI Thermal and Pan band.

Spectral Band Adjustment Factor (SBAF)

The harmonization requires adjustment of small differences due to spectral response which is specific to each instrument. In the Sen2Like processing, OLI is rescaled to S2A/MSI and S2B/MSI to S2A/MSI. Characterization results obtained with Hyperion data and proposed in the National Aeronautics and Space Administration (NASA) Harmonized Landsat Sentinel-2 (HLS) project are used [RD.2].

Data Fusion

The fusion process is the main step involved in the production of fuzzed products (L2F processing level) with Landsat 8 image pixel spacing fully aligned on Sentinel-2 image pixel spacing.

In the literature, in most cases high temporal revisit of one sensor is combined with high spatial frequency of another similar sensor in order to produce synthetic data high frequencies in time and space ([RD.17]).

Herein the context is slightly different: compared to Landsat 8, the Sentinel 2 mission (S2A / S2B) offers the best revisit time and the best spatial resolution for most spectral bands. Landsat 8 data is used to complement Sentinel 2 data.

Also, this harmonisation process improves significantly the revisit time: the theoretical number of acquisitions of this virtual constellation (95 products / year) is increased by 30 % with respect to Sentinel-2 (S2A & S2B) only acquisitions (73 products / year).

However, an additional constraint has been set on the time lineless of this data fusion process, requiring that the product of a 10.0 m Landsat 8 image solely relies on the observations made in the past. In this context, the OLI synthetic surface reflectance measurements at MSI spatial resolution are result of time based statistical prediction at pixel level. The main shortcomings are two folds,

- Correctness of the cloud shadow classification maps associated with inputs images involved in the prediction algorithms is expected

- The past data are in some cases not appropriate to fully characterize the biophysical processes (phenological stages) or even predict abrupt changes.

It is worth noting that for situation discussed in the second point, a dedicated quality assurance information has been added into the product in order to flag inconsistent measurements.

A sensor is sensing different characteristics of a landscape, basically break into large scale and small scale features.

From image processing point of view, the large scale features are regular, not necessarily uniform, with no discontinuity whilst the small scale features are associated to contour and texture. The large / small scale features are attributed to respectively low and high spatial frequency content.

The Sen2Like approach relies on this basic decomposition for improving the spatial resolution of Landsat 8 OLI data at a given date.

The Landsat 8 existing large scale features are complemented with predicted Sentinel 2 small scale features, at least small scale features not captured by Landsat 8, as described with the following equation.

$$L8^{10m} = S_{L8}^{30\rightarrow 10} + D_{S2-L8}^{10m} \quad \text{(Equation 1)}$$

Where:

- $L8^{10m}$ is the final Landsat 8 image at the Sentinel 2 spatial resolution, deconvolution from 30.0 m to 10.0 m
- $S_{L8}^{30\rightarrow 10}$ is the original Landsat 8 image resampled from 30.0 m to 10.0 m by using bilinear interpolation, it is associated to the phase of Signal,
- D_{S2-L8}^{10m} is the image of differences, differences between 30.0 m and 10.0 m spatial resolution, this information is predicted by using Sentinel 2 data, it is associated to the amplitude of signal,

For a date t, a measurement in $S_{L8}^{30\rightarrow 10}$ at image coordinates (i_0, j_0) can be expressed as an image convolution operation between the $L8^{10m}$ image and a low pass filter FS. For a small 3 x 3 window w, following mathematical relationship comes up:

$$L8_{t,i_0,j_0}^{30\rightarrow 10} = L8_{t,w}^{10m} * FS = \frac{1}{9}\sum_{S_{i,j}\in W} S_{i,j} = \frac{1}{9}\left(\sum_{\substack{S_{i,j}\in W \\ (i,j)!=(i_0,j_0)}} S_{i,j} + S_{i_0,j_0}\right) = R + \frac{1}{9}S_{i_0,j_0} \quad \text{(Equation 2)}$$

Where main quantities have been defined before. If any, it is worth noting that $S_{i,j}$ are unknown.

Similarly, for a date t, the resolution difference between the two 10.0 and 30.0 m images, corresponding to details seen by S2 and not seen by LS8, D, can be appreciated as an image convolution operation between the $L8^{10m}$ image and an high pass filter FD.

Also, for a small 3 x 3 window w, still centred at image coordinates (i_0, j_0), the following mathematical relationship comes up:

$$D_{t,i_0,j_0}^{10m} = L8_{w}^{10m} * FD$$

Where $L8_{w}^{10m}$ is the measurement window from the final Landsat 8 image. The resolution difference, D_{t,i_0,j_0}^{10m}, is unknown.

The assumption is made that this quantity can be estimated by using Sentinel 2 observations performed in the past (before the date t). In this context, the following mathematical relationship can be proposed:

$D^{10m}_{t,i_0,j_0} = f_l((S2_D)_{t-1}, \ldots, (S2_D)_{t-K},)_{i_0,j_0} + \varepsilon_{i_0,j_0}$ (Equation 3)

Where:

- $(S2_D)_{t-i} = (S2^{10m}_{t-i,w} - S2^{10m}_{t-i,w} * FS)$ with $i \in [1, K]$
- ε_{i_0,j_0} is the error term
- f_l is the best linear model prediction function minimizing error term

Also, with reference to (Eq.1, Eq.2,), the resolution difference image for the concerned scaling, can be expressed as follow:

$$D^{10m}_{t,i_0,j_0} = S_{i_0,j_0} - \frac{1}{9}\left(\sum_{\substack{S_{i,j} \in W \\ (i,j)!=(i_0,j_0)}} (S_{i_0,j_0} + S_{i,j}) \right) = \frac{8}{9} S_{i_0,j_0} - R$$

Following linearity assumption, one can adopt for simplicity,

$$S_{i_0,j_0} = \hat{S}_{i_0,j_0} + \varepsilon_{i_0,j_0}$$

It is now possible to compute the error term of the process, $\left|D^{10m}_{t,i_0,j_0} - \widehat{D^{10m}_{t,i_0,j_0}}\right|$ as follow:

$$Error_{term} = \left| (\frac{8}{9}\, \varepsilon_{i_0,j_0} - \frac{1}{9} \sum_{\substack{S_{i,j} \in W \\ (i,j)!=(i_0,j_0)}} (\epsilon_{i,j})) \right|$$

The following images have been extracted from the MGRS 31TFJ Sen2Like dataset. A side by side comparison of surface reflectance images for different landscapes (Valley, Field, Salt) shows the value added of the Sen2Like processing. Compared to 30.0 m data, the 10.0 m process enhances the image contours and enriched the image texture. Furthermore, the noise, a major drawbacks of these kind of approach is very limited.

Valley

Figure 3: Landsat 8, 10.0 m / 30.0 m side by side comparison, Valley type regions.

Figure 4: Landsat 8, 10.0 m / 30.0 m side by side comparison, Crop fields type regions.

Figure 5: Landsat 8, 10.0 m / 30.0 m side by side comparison, Salt type regions.

2.3 Processing details

The software is composed of "Thematic" blocks, one for each algorithm to be implemented, and also of "Generic" building blocks dedicated to data access, packaging, etc.

Concerning the geometric processing, Table 3 lists, depending on the band group defined in

Table 2, the applicability of each sub-processing (Framing, Co-registration, QC check). The initial MGRS tile geographical definition is always kept for subsequent processing, and then no framing is applied on Sentinel 2 data. Within the same frame, the co-registration process shifts image data depending on geolocation errors. If errors are too strong, a situation where missing pixel in margin area might be theoretically observed. Experience has shown that this situation happens only rarely. Nevertheless, it is worthy to highlight that nominally Sentinel-2 tiles already include 10km of overlap located at East and South sides.

Table 3 : Sen2Like geometric processing applicability.

Processing	Geometric Processing					
	Framing		Co-registration		QC Check	
Mission	LS8	S2	LS8	S2	LS8	S2
Coastal + SWIR	x		x	x	x	x
RGB	x		x	x	x	x
NIR	x		x	x	x	x
Pan + Thermal	x	N/A	x	N/A	x	N/A
Red Edge	N/A		N/A	x	N/A	x

As shown in table just hereafter, for some Sentinel 2, Landsat 8/9 specific bands, there is no processing applied. More inter-calibration, not listed in the table, is exclusively applied to Sentinel-2 data.

Table 4 : Sen2Like NBAR, SBAF, Fusion processing applicability.

Processing	NBAR		SBAF		Fusion	
Mission	LS8	S2	LS8	S2	LS8	
Coastal + SWIR	(1)x	(1)x	x	x	(3)x	30 m > 20 m
RGB	x	x	x	x	x	30 m > 10 m
NIR	x	x	x	(2)x	x	30 m > 20 m
Pan + Thermal	-	-	-	-	-	
Red Edge	-	-	-	-	-	

(1) - Only SWIR, (2) - Only S2 B8A, (3) - Only SWIR

As discussed more in details in this document, the Sen2Like fusion process applied to one specific product observed at a given date (d) requires also a sample of products observed prior to the date (d). With a minimum number of 2 products, process is more reliable when these products are as close as possible from the date (d).

2.4 Operational modes

Beyond operational modes, the s/w optimizes the input data selection based on fundamental criteria defined in Ground Image Processing Parameter (GIPP) file. These GIPP parameters are notably;

- The Cloud coverage parameter to filter MGRS image tile contaminates with strong nebulosity;
- The percent coverage parameter to discard products with a very limited geographic overlap with respect to MGRS tile extent;

- The priority parameter to manage the Landsat collection tiers inventory structure (Real-Time (RT), Tiers 1, Tiers 2).

This mechanism has been developed to support single tile mode and is de facto used for multi tile mode.

The processing performed in each operational mode are defined in the software configuration and can be overridden by command line arguments (see 5.1.2.1 - Processing).

Operational modes only differs in the way inputs are provided.

2.4.1 Single-Tile Mode

The processor considers as input one MGRS tile and a period (start time / end time). By using this mode, all corresponding LS8 (LS9) and S2 data are processed. An MGRS multi temporal stack (L2F / L2H) is delivered.

2.4.2 Multi-Tile Mode

The processor considers as input an AOI contained in a GeoJSON file and a period of time (start time / end time). By using this mode, all MGRS S2 tiles and LS8 (LS9) scenes that overlap the AOI within a selected time period are processed. An AOI based multi temporal stack (L2F / L2H) is delivered.

It is worth noting that there is no clipping of output data to match exactly the geographical coverage of input AOI.

2.4.3 Product Mode

The product mode allows to feed the processor directly with an input product, LS8 (LS9) or S2, and is mainly useful for debugging purpose, or for environments that provide already the single/multi tile mechanism. The production of a time series on a specific MGRS tile, like the single tile does, would then require to manually execute a run of the processor for each input product of the time series, respecting the order of the acquisition dates (as many runs as input products).

2.5 Design and Implementation

The software is developed in python (version 3) as an open source solution. The source code of Sen2Like is available in a Github repository, as part of a SNAP subfolder; https://github.com/senbox-org/Sen2Like

The software is composed of "Processing" blocks, one for each algorithm to be implemented, and also of "Generic" building blocks dedicated to data access, packaging, etc. The "Processing" blocks present a generic interface. This simplifies the integration of new blocks, but also the switch of algorithms for a same thematic

An overview of the Sen2Like tool design is shown in the figure just here after. The idea is that each thematic blocks is implemented as a python class which presents a generic interface to the main program and uses internally the specific python packages dedicated to the thematic.

For example the class "S2L_Atmcor" defines the thematic block for atmospheric correction, and is based on the "atmcor" package, which contains the smac module and other atmospheric-oriented functions.

The data access layer and the orchestration of the software is supported by several classes and modules. For example the configurations of the processing and the thematic blocks are managed in a specific class called "S2L_config".

The access to the product, its metadata and its data is managed through 4 different classes, "S2L_Product", "S2L_HLS_Product", "BaseReader" and "S2L_Image". The packaging of the output product is also implemented in tow class, called "S2L_PackagerL2F" and "S2L_PackagerL2H". The packaging of the output product metadata are implemented within "QI_MTD" module.

BaseReader and S2L_Product are generic classes that are specified to match different product types:

- BaseReader: Manage metadata access to a product type (Landsat8, Sentinel-2)
- S2L_Product; Represent a product type in the software, with a generic interface (Landsat8, Sentinel-2)

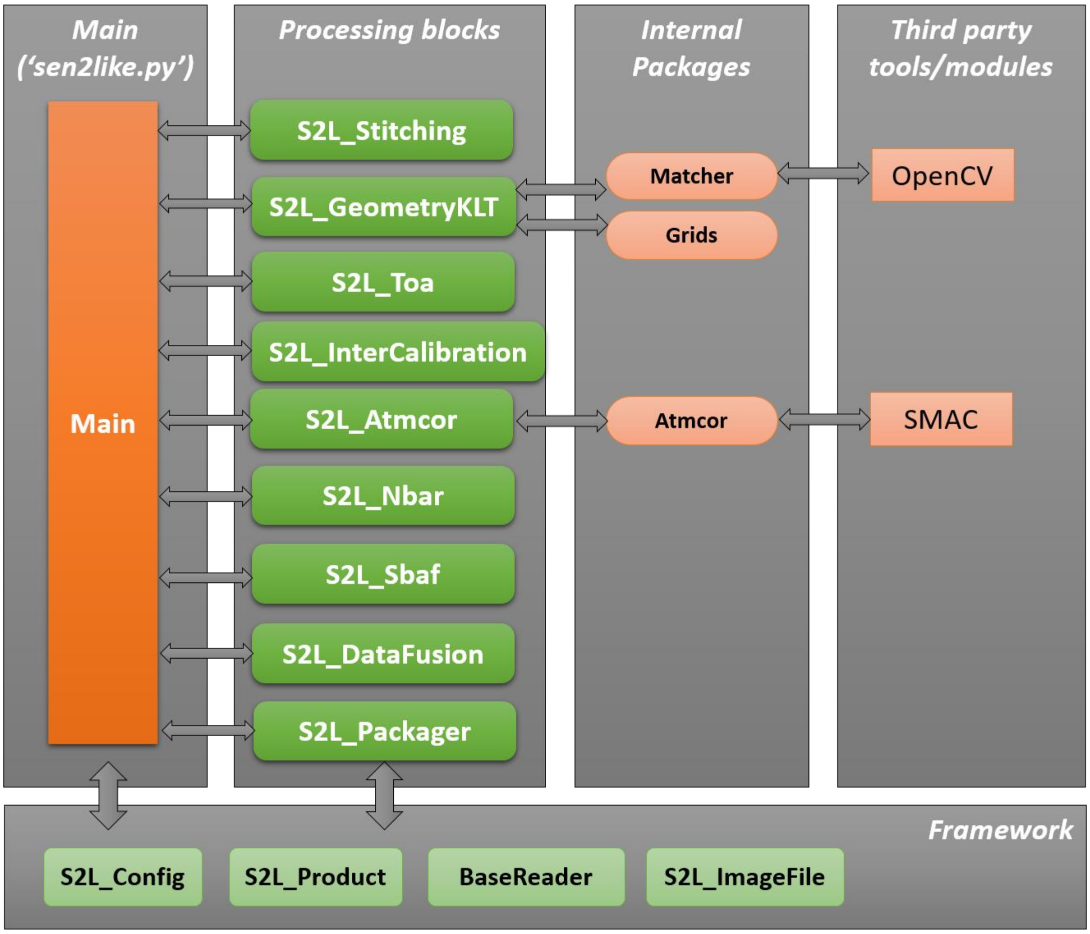

Figure 6: Overall software design

2.6 Limitations / disclaimers

The table lists the known limitations related to the usage of the Sen2Like software.

Table 5 - Limitations and disclaimers of the Sen2Like software.

Item	Description
UTM handling	In single-tile, only Landsat8 products with the same UTM projection (same zone) than the MGRS tile are considered as eligible, for subsequent processing.
Mixing Input data type : • Input S2 L2A & LS8 L1C (SMAC)	Accuracy lost due to this selection because for same thematic block different processing methods.

Item	Description
• Input S2 L2A & S2 L1C (SMAC or Sen2Cor)	Feedbacks from the S/W with a Warning is provided (TBC)
No clipping	The AOI definition (json) is not included into the L2H/L2F product. Although, Sen2Like does not perform clipping to fit exactly with the area defined by the AOI. It is up to user to do this process.
MGRS Selection based on ROI ((Improvements Foreseen))	Even if Landsat WRS does not intersect the ROI the WRS corresponding to MGRS is selected
Data selection and cloud percentage	In multi tile mode and during data selection process, a product can be discarded because of high cloud cover score. A limitation is that the cloud cover can be related to an area in the tile that is finally outside from the AOI.
Missing data (No correction foreseen)	The co-registration process shift image data depending on geolocation errors. If errors are too strong, a situation where missing pixel in margin area might be theoretically observed.
Data Fusion Algorithm	The efficiency of the data fusion algorithm strongly depends on the reliability of the quality assurance mask. For instance, even if post processing are performed the quality of the L1C cloud mask remains degraded compared to the L2A Scene Classification mask. It results in inconsistencies in the predictive scheme, as discussed above in $2.2

3. INSTALLATION

3.1 Package installation

Installation of Anaconda

```
curl  https://repo.anaconda.com/archive/Anaconda3-2020.02-Linux-
x86_64.sh --output  Anaconda3-2020.02-Linux-x86_64.sh
```

Installation of Sen2Like code

- Using git:

```
git clone https://github.com/senbox-org/Sen2Like
cd Sen2Like
```

- Or from a downloaded archive:

```
unzip Sen2Like.zip
cd Sen2Like
```

Create a conda virtual environment with required packages

```
conda create -n Sen2Like --file requirements.txt -c conda-forge
```

Activate conda virtual environment

```
conda activate Sen2Like
```

Installation of dependencies

```
sudo apt-get install mesa-libGL
```

4. RUNNING TIME

The software contains a main python file, "Sen2Like.py", which is expected to be run through a command line with arguments and options.

```
usage: Sen2Like.py [-h] [-v] [--refImage PATH] [--wd PATH] [--conf PATH]
                   [--confParams STRLIST] [--bands STRLIST]
                   [--no-run] [--intermediate-products]
                   [--parallelize-bands] [--debug]
                   [--no-log-date]

                   {product-mode,single-tile-mode,multi-tile-mode} ...
```

The main argument is the operational mode to be used:

Main Argument	Description
single-tile-mode	Run the tool on a MGRS tile. Corresponding products will be loaded.
multi-tile-mode	Run the tool on a AOI defined in a geojson. Corresponding MGRS tile will be inferred and products will be loaded. It is equivalent to run a single-tile mode for each matching tile. In multi-tile mode, multiprocessing can be used to speed-up computation time
product-mode	Run the tool on a single product

Depending on the choice of the operational mode, some options offered by the software can be different (see "Specific Options"). Other options are generic (see "Generic Options").

4.1 Specific Options

4.1.1 Single Tile Mode

Argument	Description
tile	Id of the MGRS tile to process
Options	**Description**
start-date	Beginning of period (format YYYY-MM-DD)
end-date	End of period (format YYYY-MM-DD)
l2a	Indicates if Level-2A products have to be considered. If not set, Level-1C products will be processed (default: False).

4.1.2 Multi Tile Mode

Argument	Description
roi	Geojson file containing the AOI to process
Options	**Description**
start-date	Beginning of period (format YYYY-MM-DD)
end-date	End of period (format YYYY-MM-DD)
jobs	Number of tile to process in parallel
l2a	Indicates if Level-2A products have to be considered. If not set, Level-1C products will be processed (default: False).

4.1.3 Product Mode

Argument	Description
product	Landsat8 L1 product path / or Sentinel2 L1C / L2A product path
tile	The tile on which is located the provided product.

4.2 Generic Options

Generic Options	Description
version	Display software version.
conf	Sen2Like configuration file (default: SEN2LIKE_DIR/conf/config.ini) See chapter 5 - CONFIGURATION for details.
confParams	Overload parameter values (default: None). Given as a "key=value" comma-separated list. Example: --confParams "doNbar=False,doSbaf=False" *Parameters set in the confParams command line "option" supersede the parameters in the configuration file.*
Wd	Working directory (default : /data/production/wd)

Bands	Bands to process as coma separated list (default: ALL)
refImage	Reference image (use as geometric reference) See chaper 5.2.1 for details.
no-run	Do not start process and only list products (default: False). *Usually interesting before starting a single-tile or multi-tile processing in order to evaluate and verify all the products that will be processed.*
intermediate-products	Generate intermediate products (default: False) For each processing block, each band, the intermediate output image is written saved into a file in the working directory.
Parallelize-bands	Process band in parallel
Debug	Display debug messages (default: False)
--no-log-data	Do not store timestamp in output log (default: False). Mainly used for log files comparison.

Note 1)
As shown in the product breakdown above, the process output both the 30-m dataset and the 10-m dataset. In order to get only 30-m dataset, the "doPackagerL2F" process should be disabled.

Note 2)
There are two important points for using the fusion process:

- ***It is not required to launch Sen2Like two times for one LS8/S2 product**, the Sen2Like command with "doPackagerL2F" process set to True allows to generate a product including both 10.0 m & 30.0 m bands.*
- *Prior launching Sen2Like on LS8 product with the "doPackagerL2F" option, **it is important to launch Sen2Like on at least the two S2 <u>past</u> products** observed <u>as close as possible</u> the LS8 product observation date.*

4.3 Examples

4.3.1 Single Tile Mode

```
python sen2Like.py single-tile-mode 31TFJ --wd ~/wd --refImage
/data/HLS/31TFJ/L2F_31TFJ_20170103_S2A_R008/L2F_31TFJ_20170103_S2A_R008_B0
4_10m.TIF
```

```
python sen2Like.py single-tile-mode 31TFJ --wd ~/wd --refImage
/data/HLS/31TFJ/L2F_31TFJ_20170103_S2A_R008/L2F_31TFJ_20170103_S2A_R008_B0
4_10m.TIF --conf config.ini --start-date 2017-01-01 --end-date 2017-12-01
--no-run --confParams  archive_dir=/data/S2L,coverage=0.5
```

4.3.2 Multi Tile Mode

```
python sen2Like.py multi-tile-mode roi.geojson --wd ~/wd --refImage
/data/HLS/31TFJ/L2F_31TFJ_20170103_S2A_R008/L2F_31TFJ_20170103_S2A_R008_B0
4_10m.TIF
```

4.3.3 Product Mode

```
python sen2Like.py product-mode /eodata/Sentinel-
2/MSI/L1C/2017/01/03/S2A_MSIL1C_20170103T104432_N0204_R008_T31TFJ_20170103
T104428.SAFE --wd ~/wd --tile 31TFJ --bands B04
```

5. CONFIGURATION

5.1 Processor Configuration

The configuration of the tool is done by command-line arguments and by a configuration file.

The default location of the configuration file is `S2N2LIKE_DIR/conf`.

An example is provided in Appendix chapter 0.

5.1.1 Configuration File Format

Two configuration file formats are supported:

- INI file (.ini)
- GIPP file (.xml)

5.1.2 Configuration File Sections

In the configuration file the parameters are grouped into several sections, dedicated to the configuration of the orchestration of the processing blocks (on/off), the configuration of the data archives (inputs, outputs, auxiliary), and the internal configuration of the processing blocks.

5.1.2.1 Processing

Enable or disable a processing block based on value (`True, False`):

Parameter Name	Description	Type	Range

doStitching	Run the stitching processing	Boolean	(True, False)
doGeometryKLT	Run the geometric correction processing using KLT	Boolean	(True, False)
doToa	Run the TOA conversion	Boolean	(True, False)
doInterCalibration	Run the Inter Calibration correction	Boolean	(True, False)
doAtmcor	Run the Atmospheric correction	Boolean	(True, False)
doNbar	Run Nbar correction processing	Boolean	(True, False)
doSbaf	Run the Sbaf correction processing	Boolean	(True, False)
doFusion	Run the Fusion processing	Boolean	(True, False)
doPackager	Run the packaging processing	Boolean	(True, False)
doPackagerL2H	Run the packaging processing for HLS.	Boolean	(True, False)
doPackagerL2F	Run the packaging processing for Fusion.	Boolean	(True, False)

5.1.2.2 Directories

Indicates path for special directories:

Parameter Name	Description	Type	Range
archive_dir	Where to store resulting products	Path	-
cams_dir	Where are located CAMS monthly files	Path	-
{cams_daily_dirr}	Where are located CAMS daily files	Path	-
{cams_hourly_dir}	Where are located CAMS hourly files	Path	-

{cams_climatology_dir}	Where are located CAMS climatology files	Path	-
{dem_dir}	Where are located DEM files	Path	-
scl_dir	Where are located SCL maps files	Path	-

5.1.2.3 InputProductArchive

Describes parameters for product acquisition.
By default, two methods are described:

- **local**: products are stored in local
- **creodias**: products are located using the Creodias API

Other access method can be defined by defining custom attributes, in order to use other API. To define path, custom attributes can be defined in the configuration file.

In addition these parameters are defined in the tool and can be used in brackets {}:

- mission: Landsat8 or Sentinel2
- tile: MGRS tile
- path: WRS path
- row: WRS row

Note 3)
The user should specify only one method ('local' or 'creodias'): in one run, the software cannot switch between two methods.

5.1.2.3.1 Local

Parameter Name	Description	Type	Range
coverage	Define the coverage of the product tile in the interval	Float	[0, 1] (0-100%)
base_url	Specify where the products are stored	Path	-
cloud_cover	Maximum cloud cover (%)	Int	0 - 100

url_parameters_pattern_Sentinel2	Describe storage path for Sentinel 2 products	Pattern	-
url_parameters_pattern_Landsat8	Describe storage path for Landsat 8 products	Pattern	-

Example: with the following configuration:

```
base_url = /data/PRODUCTS

cloud_cover = 11

url_parameters_pattern_Sentinel2 = {base_url}/{mission}/{tile}

url_parameters_pattern_Landsat8 = {base_url}/{mission}/{path}/{row}
```

For a Sentinel 2 product on tile 31TFJ, the software will resolved:

```
url_parameters_pattern_Sentinel2 = /data/PRODUCTS/Sentinel2/31TFJ
```

5.1.2.3.2 Creodias API

Parameter Name	Description	Type	Range
coverage	Define the coverage of the product tile in the interval	Float	[0, 1] (0-100%
base_url	Base address of the api	URL	-
cloud_cover	Maximum cloud cover (%)	Int	0 - 100
location_Landsat8	Expression specifiying Landsat 8 filter	Pattern	-
location_Sentinel2	Expression specifiying Sentinel 2 filter	Pattern	-
url_parameters_pattern	API request url. Special parameters between brackets are replaced by defined attributes	Pattern	-
thumbnail_property	Path in result json where product path is stored	Property	-
cloud_cover_property	Path in result json where cloud cover is stored	Property	-
S2_processing_level	Level of processing for considered products. Managed by software but can be specified here	Property	

Example: with the following configuration:

```
coverage = 0.5

base_url = https://finder.creodias.eu/resto/api/collections

cloud_cover = 11

location_Landsat8 = path={path}&row={row}

location_Sentinel2 = processingLevel={s2_processing_level}&productIdentifi
er=%25{tile}%25

url_parameters_pattern = {base_url}/{mission}/search.json?maxRecords=1000&
_pretty=true&cloudCover=%5B0%2C{cloud_cover}%5D&startDate={start_date}&com
pletionDate={end_date}&sortParam=startDate&sortOrder=ascending&status=all&
{location}&dataset=ESA-DATASET

thumbnail_property = properties/productIdentifier

cloud_cover_property = properties/cloudCover

gml_geometry_property = properties/gmlgeometry
```

5.1.2.4 DemDownloader

Define parameters for Copernicus DEM acquisition.

Parameter Name	Description	Type	Range
download_if_unavailable	Download requested DEM if not available locally.	Boolean	(True, False)
dem_dataset_name	Name of the DEM dataset	Pattern	-
dem_local_url	Local url of the DEM file.	Pattern	-
dem_tmp_local_url	Local url of downloaded tiles and mosaics.	Pattern	-
dem_server_url	Url of the server for DEM acquisition.	Pattern	-

5.1.2.5 Geometry

Define parameters for geometric correction.

Parameter Name	Description	Type	Range
reference_band	The reference band to use for geometric correction. Default value is B04.	String	B01 – B12
doMatchingCorrection	Apply the matching correction	Boolean	(True, False)

doAssessGeometry	Assess geometry	Pattern	(True, False)
references_map	The path to the JSON file that contains for each MGRS tile, the file path of the corresponding reference image. (See details in $5.2.1.) Can be set to None.	Filepath	-
force_geometric_correction	Force geometric correction even if the Sentinel 2 product is refined (with Sentinel 2 Global Reference Image).	Boolean	(True, False)

5.1.2.6 Atmcor

Describe the atmospheric method to be used.

Parameter Name	Description	Type	Range
use_sen2cor	Use sen2cor to do atmospheric correction if set to True. Use SMAC otherwise.	Boolean	(True, False)
sen2cor_path	Path to the sen2cor process.py file.	Path	-

5.1.2.7 Nbar

Describe the Nbar method to be used.

Parameter Name	Description	Type	Range
nbar_methode	The method to get coefficient to build constant.	String	{ROY,VJB}
vjb_coeff_matrice_dir	The path to the VJB coefficient file directory (only use by VJB method).	Path	-

5.1.2.8 Fusion

Define parameters for fusion processing.

In order to control fusion quality, a validity threshold mask of proportional difference ((l2fimage – l2himage) / l2himage) if compute for one band.

Parameter Name	Description	Type	Range
predict_method	Predict method to use (predict or composite using most recent valid pixels)	String	(predict, composite)
predict_nb_products	Number of products needed by predict method If not enough products in predict_nb_products It is the nearest product in the past that is considered	Int	1 - N
fusion_auto_check_band	Band of the validity threshold mask.	String	B01-B12
fusion_auto_check_threshold	Maximum percentage threshold to be valid in the mask.	Float	[0, 1] (0-100%

Note 4)
In the current implementation, predict_nb_product parameter value does not exceed 2.

5.1.2.9 Stitching

Define parameters for stitching processing.

Parameter Name	Description	Type	Range
reframe_margin	Margin to add during stitching reframing (in pixels)	Int	1 - N

5.1.2.10 OutputFormat

Define modifier for written image file.

Parameter Name	Description	Type	Range
gain	Gain multiplier for output image (default: 10000)	Int	1 - N
offset	Offset to add to the output image (default: 0)	Int	0 - N

output_format	Format of output images	String	(COG, GTIFF, 'JPEG2000)

5.1.2.11 COGOptions

Parameter Name	Description	Type	Range
interleave	Interleave mode	String	-
internal_tiling	Internal tiling value	Int	-
internal_overviews	Internal overviews value	Int	-
downsampling_levels_RES	Downsampling levels for resolution RES	Int List	-
downsampling_levels	Type of downsampling levels	String	-
resampling_algo_MASK	Resampling algorithm for mask	String	-
resampling_algo	Resampling algorithm	String	-
compression	Compression method	String	-
predictor	Predicator value	Int	-

5.1.2.12 JPEG2000Options

Parameter Name	Description	Type	Range
lossless_jpeg2000	Indicates if output products in jpeg2000 format are lossless.	Boolean	(True, False)

5.1.2.13 Multiprocessing

Define parameters for multiprocessing in multi-tile-mode.

Parameter Name	Description	Type	Range
number_of_process	Maximum number of processes to run in parallel	Int	1 - N

5.1.2.14 Packager

Define packaging parameters.

Parameter Name	Description	Type	Range
quicklook_jpeg_quality	Quality for outputs quicklooks in % (default: 95%)	Int	1 - 100
json_metadata	Indicates if metadata are also generated as json	Boolean	(True, False)

5.1.2.15 Runtime

This section is overridden during runtime and contains backup of computed values. Modifying this section will have no effect.

5.2 Auxiliary Data Configuration

5.2.1 HR BRDF data

As mentioned before, through configuration, it is possible to switch between two BRDF adjustment techniques (involved in the NBAR processing).

Whilst the ROY method coefficient set is hard coded (mission product classes), the per pixel HR BRDF coefficients are stored within a NetCDF file (one file per tile). As an auxiliary data, the directory path pointing to the NetCDF file directory must be indicated in the configuration file.

In the configuration file, the HR BRDF directory path is specified with the *vjb_coeff_matrice_dir parameter.*

The only constraint on the NetCDF file name is set on the end of the name and must be "_BRDFinputs.nc". An example of HR BRDF NetCDF file is as follow:

"L2F_T31UGR_2019_2019_S2_R108_BRDFinputs.nc"

The NetCDF file content is:

- **Coordinates**: geographical x, y coordinate
- **Data variables**: The V0 and V1 coefficient for each s2 bands with the key: *V0_tendency_{band name (ex: Red}*
- **Attributes**:
  - TILE
  - SPATIAL_RESOLUTION
  - BANDS : list of band name (ex: Red)
  - BANDS_NUMBER : list of band id (ex: B04)

5.2.2 Geometrical Reference Data

Sen2Like provides the capability to co-register the input products, including Sentinel-2 products. In order to enable this functionality, a reference image must be provided. The reference image shall be specified in the *references_map* file, or given within the Sen2Like command line option "--*refImage*". Note that if both are provided, the option "--*refImage*" has the priority.

The path of the *references_map* file must be specified in the configuration file (see parameter description in $0). The format of the *references_map* file is JSON, and gives for each MGRS tile identifier (e.g. "31TFJ"), the path to the reference image file to be used. An example of this file is given in Appendix C.

This reference image shall be in the geometry of the MGRS tile to process (same extent). The resolution of the reference image is not fixed, but it is recommended to provide a resolution equal to the highest resolution of the input products, i.e. 10m.

Then Sen2Like will automatically resample the reference image when necessary, for instance for Landsat8 30m bands. The resampling process is done once, and then the resulting image, saved in the reference image directory, is directly reused.

The format of the reference image is usually jpeg2000 or GeoTiff. But it can be actually any format compatible with the GDAL library.

At end, the following configuration must be verified:

- *reference_band*: the equivalent Sentinel-2 band to use for matching with the provided reference image (B04 by default)
- *doMatchingCorrection*: enable the co-registration correction
- *doAssessGeometry*: enable co-registration assessment
- *references_map:* path to the references map file

For more details on these parameters, see section 0.

5.2.3 SCL map

When processing Sentinel-2 L1C products, the corresponding Scene Classification (SCL) map available in Sentinel-2 L2A products can be used to generate the valid pixels mask associated with different Sen2like processing blocks (e.g. fusion).

Please note that this is a temporary solution for the processing of Sentinel-2 L1C product without using Sen2Cor 3.0 for atmospheric correction.

This database of L2A SCL maps needs to be populated by the user, with SCL maps in Cloud Optimized GeoTIFF format (.tif extension).

The access to this database is then configured with the *scl_dir* parameter (see 5.1.2.2).

If *scl_dir* is not set, the valid pixels mask is constructed based on Sentinel-2 L1 cloud mask.

Characteristics of the SCL database are given hereafter.

SCL database

Description:

One directory per tile (e.g. 31TFJ).
Each directory contain SCL COG files generated with Sen2Cor 2.10 with the "sc_cog" option, following the naming convention: T31TFJ_20201228T104349_SCL_60m.tif

Examples of directories:
31TFJ 12SVB

Examples of SCL COG file:
31TFJ/T31TFJ_20201228T104349_SCL_60m.tif

5.2.4 Digital elevation model

DEM search with open search interface

•http://panda.copernicus.eu/Mc3OpenSearch/webapi/Services/getProducts/?parentIdentifier=COP-DEM_GLO-90-DTED/2019_1

Automatic Download for the user

Note 4)
The DEM download is not yet integrated in the sen2like workflow, but it can be manually downloaded with the use of the new dem_downloader module.

Once the configuration file is filled with accurate server url and filename patterns, the dem can be retrieved with the following commands:

```
from dem_downloader import DemDownloader

dem_downloader = DemDownloader(r'..//conf/config.ini')
dem = dem_downloader.get(TILE, resolution=RESOLUTION,
hcs_code=EPSG_CODE)
```

Where TILE is the tile for the DEM and RESOLUTION is the resolution for the output DEM image file.

5.2.5 On the CAMS

For enabling the use of CAMS in Sen2Like, the database must be prepared.

Note 5)

The CAMS monthly database for the Year 2020 is available

http://185.178.85.51/CAMS/

When ready, the access to this database is configured with the *cams_dir* parameter (see 5.1.2.2).

It CAMS is not configured, a default AOT value will be used.

Sen2Like is able to retrieve atmospheric parameters from 4 type of databases, derived from ECMWF/CAMS:

- CAMS Monthly database (analysis)
- {CAMS Daily database (near real time)}
- {CAMS Hourly database (near real time)}
- {CAMS climatology database (climatology)}

Specificities of each database are given below.

CAMS Monthly database (analysis)

Description:

One directory per month, with a naming as 201704 for April 2017. Each directory contain a single NetCDF file with the following naming convention: CAMS_archive_aod550_tcwv_msl_gtco3_[reanalysis|analysis]_0H_6H_12H_18H_YYYY-MM.nc.

Each NetCDF file covers the whole month, with data every 6H, and contain data for the 4 parameters

Format of NetCDF file is such as provided by the ECMWF/CAMS API.

Examples of directories:

201701 201703 201705 201707 201709 201711 201801 201803 201805 201807 201809 201811 201901 201903 201905 201907 201909 201911 202001

Examples of NetCDF files:

201601/CAMS_archive_aod550_tcwv_msl_gtco3_reanalysis_0H_6H_12H_18H_2016-01.nc
201805/CAMS_archive_aod550_tcwv_msl_gtco3_analysis_0H_6H_12H_18H_2018-05.nc
201806/CAMS_archive_aod550_tcwv_msl_gtco3_analysis_0H_6H_12H_18H_2018-06.nc

{CAMS Daily database (analysis)}

Description:

One directory per day, with a naming as 20210417 for 17th of April 2021. Each directory contain a single NetCDF file with the following naming convention:CAMS_archive_aod550_tcwv_msl_gtco3_[analysis]_0H_6H_12H_18H_YYYY-MM-DD.nc

Each NetCDF file covers 4 times of the day (0h, 6h, 12h, 18h) and contain data for the 4 parameters

Format of this NetCDF file is such as provided by the ECMWF/CAMS FTP Near-Real-Time server.

Examples of directories:

20210201 20210202 20210203 20210204 20210205 20210206 20210207 20210208

Examples of NetCDF files:

20210501/CAMS_archive_aod550_tcwv_msl_gtco3_analysis_0H_6H_12H_18H_2021-05-01.nc
20210502/CAMS_archive_aod550_tcwv_msl_gtco3_analysis_0H_6H_12H_18H_2021-05-02.nc

CAMS Hourly database (near real time)

Description:

One directory per 12hours, with a naming as 2020040812 for 2020/04/08 12:00 . Each directory contain a list of NetCDF file, one per parameter and per hour, with a forecast until 12hours.

Format of this NetCDF file is such as provided by the ECMWF/CAMS FTP Near-Real-Time server.

Examples of directories:

2020040700 2020040712 2020040800 2020040812 2020040900 2020040912 2020041000 2020050300

Examples of files (for 1 directory):

2020040912/z_cams_c_ecmf_20200409120000_prod_fc_sfc_001_aod550.nc
2020040912/z_cams_c_ecmf_20200409120000_prod_fc_sfc_001_gtco3.nc
2020040912/z_cams_c_ecmf_20200409120000_prod_fc_sfc_002_aod550.nc
2020040912/z_cams_c_ecmf_20200409120000_prod_fc_sfc_002_gtco3.nc
2020040912/z_cams_c_ecmf_20200409120000_prod_fc_sfc_003_aod550.nc
2020040912/z_cams_c_ecmf_20200409120000_prod_fc_sfc_003_gtco3.nc
…
2020040912/z_cams_c_ecmf_20200409120000_prod_fc_sfc_010_aod550.nc
2020040912/z_cams_c_ecmf_20200409120000_prod_fc_sfc_010_gtco3.nc
2020040912/z_cams_c_ecmf_20200409120000_prod_fc_sfc_011_aod550.nc
2020040912/z_cams_c_ecmf_20200409120000_prod_fc_sfc_011_gtco3.nc
2020040912/z_cams_c_ecmf_20200409120000_prod_fc_sfc_012_aod550.nc
2020040912/z_cams_c_ecmf_20200409120000_prod_fc_sfc_012_gtco3.nc

CAMS Climatology database (climatology)

Description:

Climatology files are generated manually from CAMS archive. For each parameter, they present one value per day of year. The role of the climatology database is about having a backup strategy, i.e. if any valid data cannot be found in other databases.

Data file format:

The format is a TIF internal format. Size and extent of the data corresponds to official netcdf data. File are single date.

Example of files:

CAMS_Climatology_2010-2019_msl_DOY_355.tif

CAMS_Climatology_2010-2019_tcwv_DOY_001.tif}

5.3 Processing Blocks Configuration

This chapter gives information that might be useful for integrating new processing blocks.

The orchestration of the chain of blocks is at the level of the main program ("Sen2Like.py" file). It starts with the ingestion of the input product, triggering metadata and product information extraction. When the thematic blocks are eligible to a "**band by band" process**, the orchestrator is able to loop the processing to cover all the requested band, the band set being specified and by default it is all the band in the product that are processed.

The 'band by band' approach has been developed to ease parallelization in the deployment.

`[It is not all process that are eligible to band by band process, as for instance the data fusion, the mask creation]`

The "band set" approach has been set up in order to ensure quick processing of expected band.

The design has considered two type of configuration:

- An internal configuration to allow the specification of a module in term of parameter and applicability
- An external configuration to allow orchestration of the different modules.

The list of the thematic blocks to be enabled for the processing is dynamically set up from the external configuration (ON/OFF switches) and the internal configuration.

The internal configuration lists the names of the available classes that implement thematic blocks. In addition they are some parameters, like the applicability for LS8/S2.

In the external configuration, for each thematic block, the user can add an ON/OFF switch. The name of the parameter is the name of the class, but where the prefix 'S2L_' is replaced by 'do'. For example the ON/OFF parameter of the atmospheric correction block, implemented through the 'S2L_Atmcor' class, is: 'doAtmcor'.

Table 6: Example of internal configuration (declaration of building blocks)

```
PROC_BLOCKS['S2L_Stitching'] = {'extension': '_STITCHED.TIF', 'applicability': 'L8_L9_S2'}

PROC_BLOCKS['S2L_GeometryKLT'] = {'extension': '_REFRAMED.TIF', 'applicability': 'L8_L9_S2'}

PROC_BLOCKS['S2L_Toa'] = {'extension': '_TOA.TIF', 'applicability': 'L8_L9_S2'}

PROC_BLOCKS['S2L_Atmcor'] = {'extension': '_SURF.TIF', 'applicability': 'L8_L9_S2'}

PROC_BLOCKS['S2L_Nbar'] = {'extension': '_BRDF.TIF', 'applicability': 'L8_L9_S2'}

PROC_BLOCKS['S2L_Sbaf'] = {'extension': '_SBAF.TIF', 'applicability': 'L8_L9_S2'}

PROC_BLOCKS['S2L_PackagerL2H'] = {'extension': None, 'applicability': 'L8_L9_S2'}

PROC_BLOCKS['S2L_Fusion'] = {'extension': '_FUSION.TIF', 'applicability': 'L8_L9'}

PROC_BLOCKS['S2L_Packager'] = {'extension': None, 'applicability': 'L8_L9_S2'}

PROC_BLOCKS['S2L_PackagerL2F'] = {'extension': None, 'applicability': 'L8_L9_S2'}
```

APPENDIX B: EXAMPLE OF CONFIGURATION FILE

```
[Processing]
doStitching = True
doGeometryKLT = True
doToa = True
doAtmcor = True
doNbar = True
doSbaf = True
doFusion = True
doPackager = False
doPackagerL2H = True
doPackagerL2F = True

[Directories]
archive_dir = /data/HLS

cams_dir = /data/CAMS/monthly
cams_daily_dir = /data/CAMS/daily
cams_hourly_dir = /data/CAMS/hourly
cams_climatology_dir = /data/CAMS/climatology/v1

dem_dir = /data/DEM
```

```
[InputProductArchive]
coverage = 0.5
# Local
base_url = /data/PRODUCTS
url_parameters_pattern_Sentinel2 = {base_url}/{mission}/{tile}
url_parameters_pattern_Landsat8 = {base_url}/{mission}/{path}/{row}

# Creodias
;base_url = https://finder.creodias.eu/resto/api/collections
;cloud_cover = 11
;location_Landsat8 = path={path}&row={row}
;location_Sentinel2                                                       =
processingLevel={s2_processing_level}&productIdentifier=%25{tile}%25
;url_parameters_pattern                                                   =
{base_url}/{mission}/search.json?maxRecords=1000&_pretty=true&cloudCover=%
5B0%2C{cloud_cover}%5D&startDate={start_date}&completionDate={end_date}&so
rtParam=startDate&sortOrder=ascending&status=all&{location}&dataset=ESA-
DATASET
;thumbnail_property = properties/productIdentifier
;cloud_cover_property = properties/cloudCover
;gml_geometry_property = properties/gmlgeometry

[DemDownloader]
download_if_unavailable = True

# Local storage
dem_dataset_name = COP-DEM_GLO-90-DGED__2019_1
dem_local_url                                                             =
/data/DEM/{mgrs_tile}/Copernicus_DSM_{resolution}m_{mgrs_tile}.TIF
dem_tmp_local_url = /data/DEM/{mgrs_tile}_tiles/{dem_product_name}_DEM.TIF

# Copernicus server
;tile_format = TIFF
;tile_size = 5
archive_format = tar
dem_product_name = Copernicus_DSM_{arcsec:02}_{latitude}_00_{longitude}_00
#dem_server_url            =            https://cdsdata.copernicus.eu/pd-
desk/prismDownload/{dem_dataset_name}/{dem_product_name}.{archive_format}
dem_server_url                                                            =
http://172.30.16.191/DEM/{dem_product_name}.{archive_format}


[Geometry]
reference_band = B04
doMatchingCorrection = True
doAssessGeometry = B04
references_map = /data/References/references_map.json
force_geometric_correction = False
```

```
[Atmcor]
use_sen2cor = False
sen2cor_path = ../sen2cor/process.py

[fusion]
# predict_method: predict or composite (most recent valid pixels)
predict_method = predict
predict_nb_products = 2

[Stitching]
reframe_margin = 50

[OutputFormat]
gain = 10000
offset = 0
output_format = COG

[COGoptions]
interleave = PIXEL
internal_tiling = 1024
internal_overviews = 1024
downsampling_levels_10 = 2 6 12 36
downsampling_levels_15 = 2 4 8 24
downsampling_levels_20 = 3 6 18
downsampling_levels_30 = 2 4 12
downsampling_levels_60 = 2 6

downsampling_levels = variable
resampling_algo_MASK = MODE
resampling_algo = AVERAGE
compression = LZW
predictor = 1

[JPEG2000options]
lossless_jpeg2000 = True

[Multiprocessing]
number_of_process = 5

[Packager]
quicklook_jpeg_quality = 75
json_metadata = True
```

APPENDIX C: EXAMPLE OF REFERENCES MAP FILE


```
{
  "30SWJ":
"/data/References/GRI/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000094_T30SWJ_N01.01/IMG_DATA/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000094_T30SWJ_B04.jp2",
  "30TXQ":
"/data/References/30TXQ/L2F_30TXQ_20190822_S2B_R094_B04_10m.TIF",
  "32TNS":
"/data/References/GRI/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000065_T32TNS_N01.01/IMG_DATA/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000065_T32TNS_B04.jp2",
  "32TMR":
"/data/References/GRI/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000065_T32TMR_N01.01/IMG_DATA/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000065_T32TMR_B04.jp2",
  "36MXE":
"/data/References/GRI/S2A_OPER_MSI_L1C_TL_SGS__20160217T115519_A003421_T36MXE_N02.01/IMG_DATA/S2A_OPER_MSI_L1C_TL_SGS__20160217T115519_A003421_T36MXE_B04.jp2",
  "34RGS":
"/data/References/GRI/S2A_OPER_MSI_L1C_TL_MTI__20180617T111214_A015591_T34RGS_N02.06/IMG_DATA/S2A_OPER_MSI_L1C_TL_MTI__20180617T111214_A015591_T34RGS_B04.jp2",
  "20MRB":
"/data/References/GRI/S2A_OPER_MSI_L1C_TL_SGS__20160806T192619_A005870_T20MRB_N02.04/IMG_DATA/S2A_OPER_MSI_L1C_TL_SGS__20160806T192619_A005870_T20MRB_B04.jp2",
  "12SVB":
"/data/References/GRI/S2A_OPER_MSI_L1C_TL_SGS__20160420T214215_A004328_T12SVB_N02.01/IMG_DATA/S2A_OPER_MSI_L1C_TL_SGS__20160420T214215_A004328_T12SVB_B04.jp2",
  "32TQM":
"/data/References/GRI/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000122_T32TQM_N01.01/IMG_DATA/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000122_T32TQM_B04.jp2",
  "35WMQ":
"/data/References/GRI/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000022_T35WMQ_N01.01/IMG_DATA/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000022_T35WMQ_B04.jp2",
  "31TCJ":
"/data/References/GRI/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000051_T31TCJ_N01.01/IMG_DATA/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000051_T31TCJ_B04.jp2",
  "31TFJ":
"/data/References/GRI/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000008_T31TFJ_N01.01/IMG_DATA/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000008_T31TFJ_B04.jp2",
  "34TCR":
"/data/References/GRI/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000036_T34TCR_N01.01/IMG_DATA/S2A_OPER_MSI_L1C_TL_MPS__20161018T120000_A000036_T34TCR_B04.jp2"
}
```