Copernicus Space Component Sentinel Optical Mission

Performance Cluster

Sen2Water Software Design Document

Ref.: OMPC.BC.TN.010

Issue: 1.0

Date: 09/10/2023

Contract: 4000136252/21/I-BG

Customer:	ESA	**Document Ref.:**	OMPC.BC.TN.010
Contract No.:	4000136252/21/I-BG	**Date:**	09/10/2023
WP No.:	N/A	**Issue:**	1.0

Project:	Copernicus Space Component Sentinel Optical Mission Performance Cluster		
Title:	Minutes of Sen2Water Software Design Document held on 21/08/2023 by		
Author(s):	M. Boettcher, C. Brockmann, K. Stelzer, …		
Approved by:		**Authorized by**	J. Bruniquel, OPT-MPC Service Manager
Distribution:	ESA		
Accepted by ESA	Georgia Doxani Valentina Boccia		
Filename	OMPC.BC.TN.010 - i1r0 - Sen2Water SDD.docx		

ACRI-ST
260 route du Pin Montard
06904 Sophia-Antipolis, France
Tel: +33 (0)4 92 96 75 00 Fax: +33 (0)4 92 96 71 17
www.acri-st.fr

Funded by the EU and ESA

European Union

The views expressed herein can in no way be taken to reflect the official opinion of the European Space Agency or the European Union.

Changes Log

Version	Date	Changes
1.0	09/10/2023	First version

Table of content

1 Introduction

1.1 Purpose and scope

This software design document describes the functional composition of the combined Sentinel-2 Level 2 processor and its main constituents Sen2Cor and Sen2Water. It also describes the design of the software structures with directories of the runtime structure and the installation packages.

Related documents identify requirements for Sen2Water [MSI L2 RD 2023] and identify and define the inputs and outputs of the different functional elements [MSI L2 IODD 2023]. [MSI L2 ATBD 2023] defines the algorithms of the functional elements. The combined Level 2A data product is initially defined in [MSI L2 IODD 2023]. It will finally be moved to the Sentinel-2 Level 2A Product Specification Document when its update is due in the course of the Sen2Water project.

1.2 References

The following documents are referenced in this document.

Document ID	Description	Version
S2-PDGS-MPC-L2A-ATBD 2021	Level-2A Algorithm Theoretical Basis Document, S2-PDGS-MPC-ATBD-L2A, Optical Mission Performance Centre, ESA, November 2021	2.10
MSI L2W RD 2023	Sen2Water Requirements Document, OMPC.BC.RD-MSI-L2, Optical Mission Performance Centre, ESA, October 2023	1.0
MSI L2 IODD 2023	Sen2Cor and Sen2Water Input Output Data Definition, OMPC.TPGZ-BC.IODD-MSI-L2, Optical Mission Performance Centre, ESA, October 2023	1.0
OMPC.TPZG.IOD.001 2022	Sen2Cor 2.11.00 Input Output Data Definition, Optical Mission Performance Centre, ESA, November 2023	2.11
MSI L2 ATBD 2023	Sen2Cor and Sen2Water Algorithm Theoretical Basis Document, OMPC.TPGZ-BC.ATBD-MSI-L2, Optical Mission Performance Centre, ESA, October 2023	1.0
S2-PSD	Sentinel-2 Products Specification Document, *S2-PDGS-TAS-DI-PSD,* ESA	14.9
S2-PDGS-MPC-L2A-PFS	Sentinel-2 MSI – Product Format Specification	14.9
S2-PDD	GMES Space Component – Sentinel-2 Payload Data Ground Segment (PDGS), Product Definition Document	2.3
OMPC-TPZG-SUM-001	Sentinel-2 MSI – Level 2A Prototype Processor Installation and User Manual, Telespazio, OPMC, ESA	2.11
OMPC.TPZG.SRN.003	Sen2Cor 2.11.00 Software Release Note. Telespazio, OMPC, ESA, 2022	2.11

1.3 Acronyms

The following acronyms are used within this document:

Acronym	Description
ATBD	Algorithm Theoretical Basis Document
C2RCC	Case-2 Regional Coast Colour
CAMS	Copernicus Atmospheric Monitoring Service
CGLOPS	Copernicus Global Land Service
CMEMS	Copernicus Marine Service
ESA	European Space Agency
HR-OC	High Resolution Ocean Colour
IODD	Input Output Data Definition
L1C	Level 1C
L2	Level 2
L2A	Level 2A
L2W	Level 2 Water
MSI	MultiSpectral Instrument
OMPC	Optical Mission Performance Centre
PDGS	Payload Data Ground Segment
RD	Requirements Document
S2W	Sen2Water
SAFE	Standard Archive Format for Europe
SDD	Software Design Document
SNAP	Sentinel Application Platform (and Sentinel Toolboxes)

1.4 Document overview

After this formal introduction

Section 2 describes the context and main concepts of the combined Sentinel-2 Level 2 processor based on Sen2Cor and Sen2Water.

Section 3 describes functional elements of Sen2Cor, of Sen2Water, and of the combined Sentinel-2 Level 2 processor.

Section 4 describes runtime structures and delivery package artefacts of Sen2Cor, Sen2Water, and of the combined Sentinel-2 Level 2 processor.

Section 5 describes Sen2Water stand-alone, mainly by reference to previously described configurations.

2 Overview

This chapter describes the context and main concepts of the combined Sentinel-2 Level 2 processor based on Sen2Cor and Sen2Water.

2.1 Sen2Cor heritage

Sen2Cor is a processor for Sentinel-2 Level 2A product generation and formatting; it performs the atmospheric-, terrain correction of Top-Of-Atmosphere Level 1C input data. Sen2Cor creates Bottom-Of-Atmosphere, i.e. surface reflectance images; and additionally, Aerosol Optical Thickness-, Water Vapor-, Scene Classification Map and Quality Indicators for cloud and snow probabilities. Its output product format is equivalent to the Level 1C User Product: JPEG 2000 images, with three different resolutions, 60, 20 and 10 m.

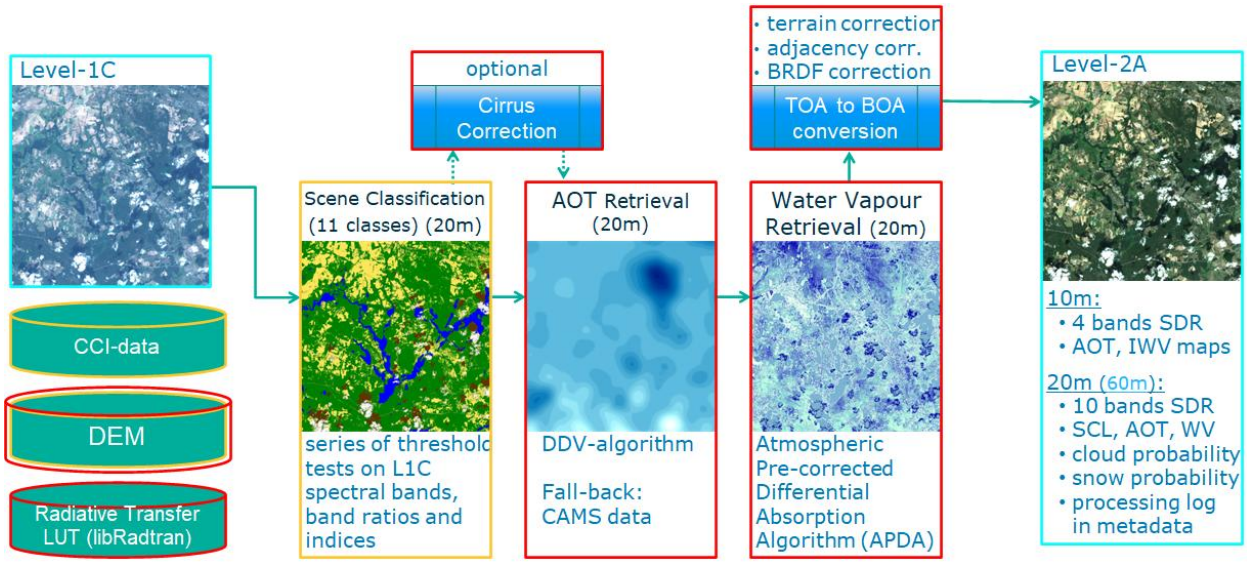

Figure 2-1: Overview of Sen2cor processing chain

Sen2Cor is designed to perform atmospheric correction of land surfaces. It is used operationally since December 2018 to generate official L2A products. Since January 2022, Sen2Cor can use aerosol information from auxiliary CAMS (Copernicus Atmosphere Monitoring Service) when it is not possible to retrieve aerosol content from the scene itself.

2.2 Sen2Water and Copernicus water quality services

The Sen2Water processor combines the pre-processing of two complementary services, the HR-OC coastal ocean processing for CMEMS and the inland water processing for CGLOPS. The two services use different chains for pre-processing that are described in two subsections. Sen2Cor operationally generates the former L2 data product and will be part of the target processing chain as well. Sen2Cor and the target chain are described in the next two subsections. The output data product of Sen2Water and the combined chain are introduced in the last subsection.

The HR-OC processing chain generates water quality parameters turbidity, suspended particulate matter chlorophyll, and a few other parameters from Sentinel-2 MSI L1C inputs. It does so for all European coastal ocean waters in a 20 km stripe from the coastline. The pre-processing chain combines C2RCC for clear water and ACOLITE for turbid water with blending in transition zones.

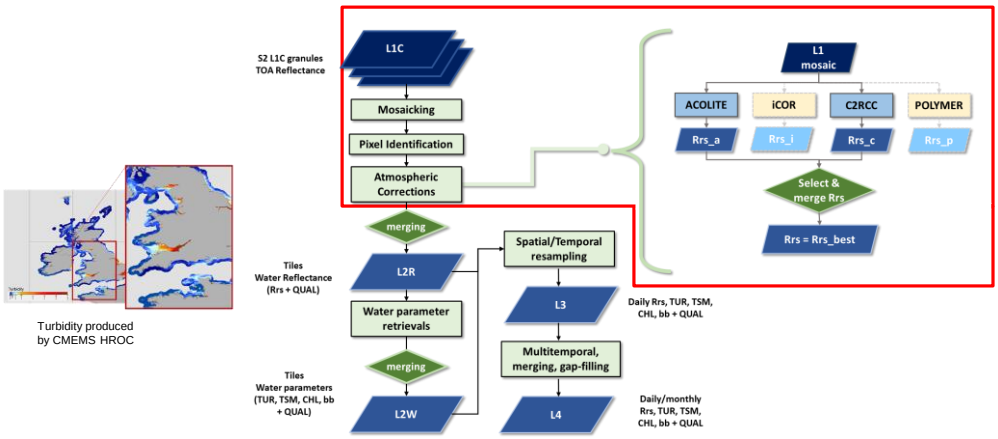

Figure 2-2: CMEMS HR-OC pre-processing chain with S2Resampling, Idepix, C2RCC, ACOLITE

In addition to C2RCC and ACOLITE, SNAP S2Resampling generates 60 m inputs with carefully determined observation angles. Granules are combined into larger mosaics before processing. A static mask is used to distinguish coastal water, coastal land, and other areas. The SNAP Idepix processor identifies cloud, ice, and dynamically re-classifies pixels as water or non-water in the coastal zones. Blending generates the AC output, a corrected mask, and flags.

The inland water processing chain generates water quality parameters turbidity, total suspended matter, chlolophyll-a, and several other parameters from Sentinel-2 MSI L1C inputs. It does so for a set of currently 225 granules distributed over Europe and Africa. The pre-processing chain uses POLYMER for atmospheric correction.

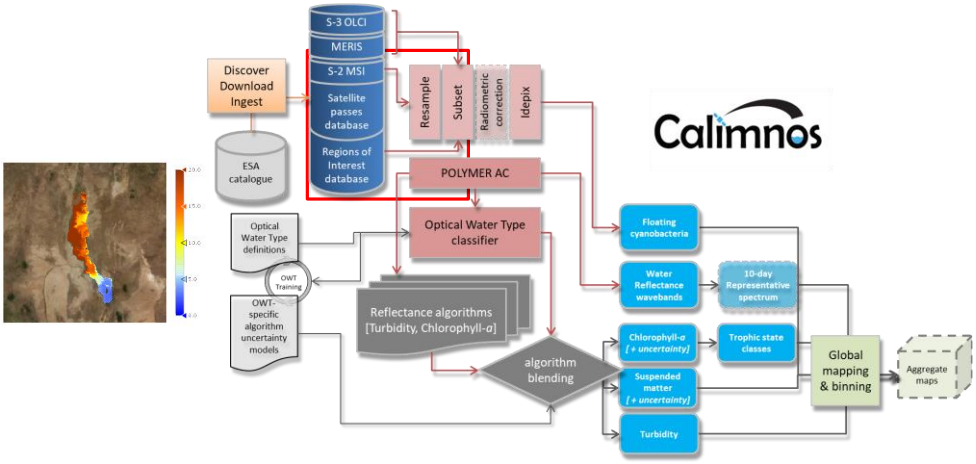

Figure 2-3: CGLOPS inland water pre-processing chain with S2Resampling, Idepix, POLYMER

Before running POLYMER, SNAP S2Resampling generates 60m inputs with observation angles. The SNAP Idepix processor distinguishes clear water dynamically from cloud, and other pixel classes.

2.3 Target processing chain configuration

There are two target configurations: Sen2Water stand-alone and the new Level 2A processor. Sen2Cor exists as stand-alone module as well. the combined Level 2A processor configuration is shown in Figure 2-4. The combined processor processes L1C in PDI format.

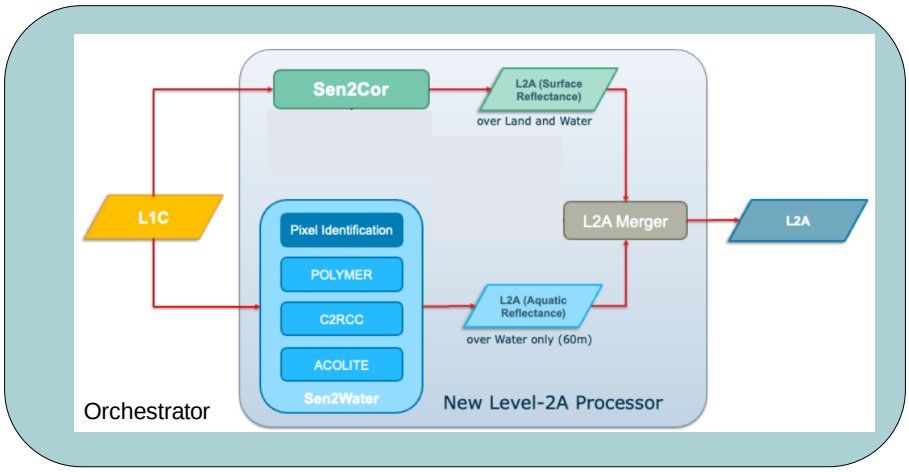

Figure 2-4: New Level 2A processor with Sen2Cor and Sen2Water

Sen2Water in stand-alone configuration processes L1C data in end user product format (EUP). The Sen2Water processor comprises S2Resampling, Idepix and three different AC processors that are applied to water areas. Selection and blending combine their outputs into the L2W aquatic reflectance data product.

The new Level-2A processor combines Sen2Cor and Sen2Water. They can be run concurrently. A merger creates the final output that includes both the generic and aquatic reflectances and respective flags.

2.4 Target output data product

Sen2Water as stand-alone data processor generates an output data product L2W in 60 m with atmospherically corrected aquatic reflectances for water areas. Corrected bands are B1, B2, B3, B4, B5, B6, B7, and B8A. It will cover ocean and inland water bodies. L2W is provided in Sentinel-2 granules of 1830 x 1830 pixels. Additional flags further characterize the quality of the retrieval. L2W is provided as NetCDF4 data file with CF-compliant metadata.

The combined processor generates the Sen2Cor output enriched by the L2W, i.e. with aquatic reflectances and flags for water areas in 60 m in addition to the outputs of Sen2Cor. The L2W content is part of the SAFE product format.

3 Elements and interfaces of the Level 2 data processors [BC, Telespazio]

This section describes functional elements of Sen2Cor, of Sen2Water, and of the combined Sentinel-2 Level 2 processor.

3.1 Elements of Sen2Cor

3.1.1 Introduction

Sen2Cor 2.10 is composed of 16 Python L2A_modules listed in Table 3-1. The Python environment is set during the installation and contains all the necessary packages for running Sen2Cor 2.10. In the following sections, a brief summary of the content and scope of the Sen2Cor 2.10 L2A_modules is reported.

Table 3-1: List of Python modules for Sen2Cor 2.10.

n	Name
1	L2A_Process
2	L2A_ProcessTilePdgs
3	L2A_ProcessDataStrip
4	L2A_ProcessTileToolbox
5	L2A_Config
6	L2A_Tables
7	L2A_SceneClass
8	L2A_SceneClass_evolution
9	L2A_KLT_Tracker
10	L2A_AtmCorr
11	L2A_Quality
12	L2A_XmlParser
13	L2A_Manifest
14	L2A_Cogconverter
15	L2A_Library
16	L2A_logger

3.1.2 L2A_Process

Main processor module. It determines, given the user's input, the processing type (PDGS or Toolbox), coordinates the calls to and from the other modules and it oversees the post-processing. The L2A_Process module keeps references of the main modules (Config, Tables, SceneClass, AtmCorr) and drives their interactions (see Figure 3-1). It also determines the sequential steps for the processing of the different products according to the user's requests and settings.

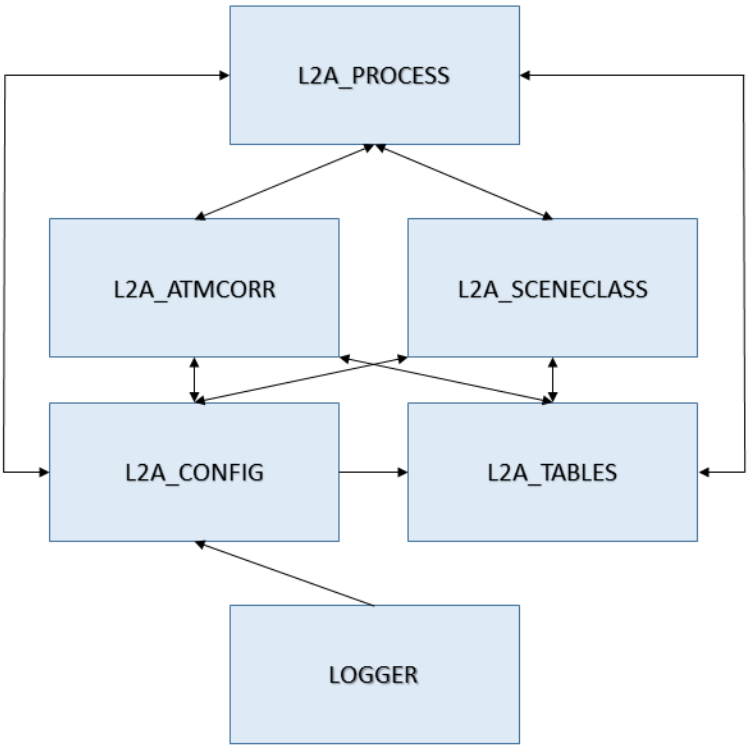

Figure 3-1: Overview of the main processor's architecture.

To summarise, the L2A_Process:

- Imports the necessary modules;
- Distinguishes between the Toolbox mode and the PDGS modes (Process Tile, Process Datastrip);
- Reads the `kwargs` passed by the users (via command lines);
- Performs a sanity check on the L1C data;
- Creates/updates the L2A Products' necessary directories and metadata;
- Coordinates the pre-, main- and post-processing.

3.1.3 L2A_ProcessTilePdgs

Main module for processing L1C tile products in PDGS mode.

3.1.4 L2A_ProcessDatastrip

Main module for processing L1C Datastrip towards L2A Datastrip. The output of this module is then used as one of the inputs for processing the L1C tile product in PDGS mode.

3.1.5 L2A_ProcessTileToolbox

Main module for processing L1C product in the standalone (Toolbox) version of Sen2Cor 2.10.

3.1.6 L2A_Config

This module provides the interface between the L1C product data (and associated metadata) and their processing and porting toward the building-up of the L2A product (and associated metadata).

3.1.7 L2A_Tables

This module contains all the required functions to perform the importing of L1C bands data, the exporting of L2A bands data and the calculations of all the auxiliary parameters/bands/quantities that are necessary to perform the L2A products processing.

3.1.8 L2A_SceneClass

This module contains all the necessary algorithms and functions to perform the scene classification of the L1C tile. This module contains the default algorithms for the scene classification until Sen2Cor 2.9. These algorithms can be still called by the optional command line --sc_classic.

3.1.9 L2A_SceneClass_evolution

This module contains all the necessary algorithms and functions to perform the scene classification evolution of the L1C tile. In this document, it will be simply referred as the Scene Classification because the Scene Classification Evolution algorithms now constitute the default selection for Sen2cor 2.10.

3.1.10 L2A_KTL_Tracker

This module contains auxiliary functions associated with the computation of the Scene Classification.

3.1.11 L2A_AtmCorr

This module contains all the necessary algorithms and functions to process the input from top of atmosphere (TOA) toward the computation of the bottom of atmosphere (BOA) representation. It, thus, performs the atmospheric correction of the L1C tile.

3.1.12 L2A_Quality

This module contains all the necessary functions to produce the L2A_Quality report (for products starting from PSD 14.9.

3.1.13 L2A_Xml_Parser

This module contains all the necessary functions for the validation of the L1C and L2A metadata.

3.1.14 L2A_Manifest

This module contains additional functions for the validation and porting of the L1C and L2A products metadata within the manifest.safe file.

3.1.15 L2A_CogConverter

(Deprecated). This module contains the functions to export COG files either as COG or JPEG2000 format. It is implemented for optionally setting the output type of the Scene Classification mask

3.1.16 L2A_Library

This module contains auxiliary functions for statistics and debugging.

3.1.17 L2A_Logger

This module handles the production of the associated logger that is produced during the processing of the L1C product.

3.2 Elements of Sen2Water

Figure 3-2 shows the elements of the Sen2Water processor.

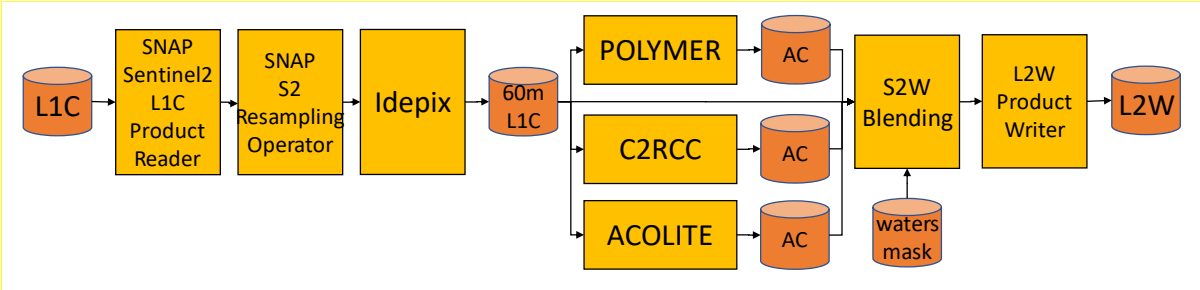

Figure 3-2: Elements of Sen2Water with Idepix, POLYMER, C2RCC, ACOLITE, and S2W Blending

The elements are modules for data transformations and storage for inputs, intermediates and the output. They are explained in the following subsections.

3.2.1 Input data reader (SNAP)

The Sentinel-2 MSI L1C data product is opened and read by the SNAP Sentinel2L1CProductReader. It is one of the readers of the Sentinel-2 Toolbox. It converts the file representation of the data product into an in-memory data organisation using the SNAP product model. This includes lazy reading, tile cache management, and metadata handling. After this step the L1C data product is available to SNAP operators in memory.

3.2.2 S2Resampling (SNAP)

S2Resampling performs resampling of reflectance bands to 60 m and a detector-aware resampling of viewing angles per pixel. It is a SNAP operator available in the Sentinel-2 Toolbox. After this step the resampled product is available to SNAP operators in memory. Evaluation is lazy and uses tile cache management as with the reader.

3.2.3 Idepix processor (SNAP)

Idepix performs pixel classification. It is a SNAP operator available as a plug-in to the Sentinel Toolbox. Idepix passes the input bands to the output. After this step the product with an additional band pixel_classif_flags is available to SNAP operators and writers in memory.

A standard SNAP writer for NetCDF4 is used to write the intermediate data product to a local storage location (working directory) of the processor run.

3.2.4 C2RCC AC processor (SNAP)

C2RCC performs atmospheric correction of water pixels. It is available as a SNAP operator in the Sentinel Toolbox.

A standard SNAP writer for NetCDF4 is used to write the intermediate data product with the water-leaving reflectances to a local storage location (working directory) of the processor run.

3.2.5 ACOLITE AC processor (Python)

ACOLITE performs atmospheric correction of water pixels. It is available as a stand-alone Python implementation. It writes water-leaving reflectances to a NetCDF4 file in a local storage location (working directory) of the processor run.

3.2.6 POLYMER AC processor (Python)

POLYMER performs atmospheric correction of water pixels. It is available as a stand-alone Python implementation. It writes water-leaving reflectances to a NetCDF4 file in a local storage location (working directory) of the processor run.

3.2.7 Blending and quality masks processor (Python)

The Sen2Water Blending processor reads the outputs of Idepix and the three AC processors from local storage (working directory) of the processor run. It blends results of C2RCC and ACOLITE in ocean water according to the algorithm used in the HR-OC (CMEMS) processing chain. It further selects the POLYMER result for inland water pixels and the blended result for ocean water pixels. In the transition zone at river mouths, it applies another blending to avoid steps in the result.

The Sen2Water Blending processor combines the masks of Idepix and the different AC processors into a set of masks for the output.

The in-memory result of Blending is formatted into the L2W data product with water-leaving reflectances, flag bands, and metadata. It is formatted in NetCDF4.

3.2.8 Processing workflow control

The processing chain is controlled by a shell script with sub-process control. The script calls the chain in sequence, but the three AC processors concurrently.

If one of the steps fails then the Sen2Water processor fails. Accepted reasons for failure are incomplete or broken inputs or misconfiguration in user-provided parameters or infrastructure-related failures like a non-writable working directory.

If the water mask does not contain any ocean water pixels or transition zone pixels then C2RCC and ACOLITE are not called. If the water mask does not contain land/inland water or a transition zone between ocean and inland water then POLYMER is not called.

3.3 Elements of the combined Level 2 processor

3.3.1 The Level 2 merger

The merger takes the outputs of the Sen2water processing and insert its contents into the related L2A product generated by Sen2Cor, according to the updated L2A product format specification referenced in section 0.

3.3.2 Processing workflow control

The processing chain is controlled by a shell script with sub-process control. The script concurrently calls Sen2Cor and Sen2Water, waits for their termination, and calls the merger to generate the L2A product.

A parameter determines whether failure of Sen2Water shall mean failure of the complete processor or shall be tolerated and an output without water pixels shall be generated. Note that there are no failures expected for Sen2Water except for external reasons like broken input product or infrastructure failures. Therefore, it is not recommended to tolerate failure.

Sen2Water does not need to be called if there are no water pixels expected in the granule, .e.g. desert areas. This information is part of the static mask used with Sen2Water. A list of granules that may contain water is compiled from the set of masks and used in this processor.

4 Software structures

This section describes structures of Sen2Cor, Sen2Water, and of the combined Sentinel-2 Level 2 processor. The runtime structure with the layout of directories after deployment defines where modules of the processor are installed and can be used at runtime. The delivery package defines the artefact that is provided for installation.

4.1 Sen2Cor software structures

4.1.1 Software runtime structure

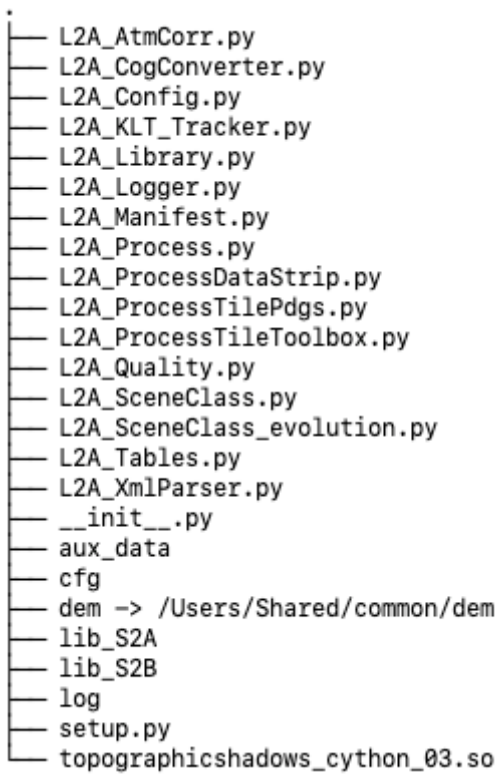

Figure 4-1: software runtime structure of a full Sen2Cor Installation

Figure 4-1 shows the software runtime structure of a full Sen2Cor Installation.

- L2A_module_name.py are the python files (described in Section 3.1 of the SDD)
- Aux_data: folder containing the auxiliary files (ESA CCI files, Global Snow Map Copernicus Missing Tiles lists at 30 and 90 m (can be a symbolic link)
- Cfg: folder containing the L2A GIPPs and associated schemas, safe directories and schemas for the available PSDs, L2A_Quality xml file and associated schema, other configuration files.
- Dem: directory where downloaded DEM are stored (can be a symbolic link as shown in Figure 4-1).
- Lib_S2A, Lib_S2B: folders containing the Look-Up-Tables for Copernicus Sentinel 2-A and 2-B
- Log: folder containing the resulting log files from processes.
- Topographicshadows_cython_03.so: library dedicated to the computation of the topographic shadow for the Scene Classification Evolution.

The software can be deployed using the provided installation package on the ESA Step Page. Detailed Instructions can be found in the associated SUM and SRN.

4.2 Sen2Water software structures

This chapter describes the integration of the processor elements of Sen2Water on the structural level.

4.2.1 Software runtime structure

Figure 4-2 shows the software runtime structure of a Sen2Water installation.

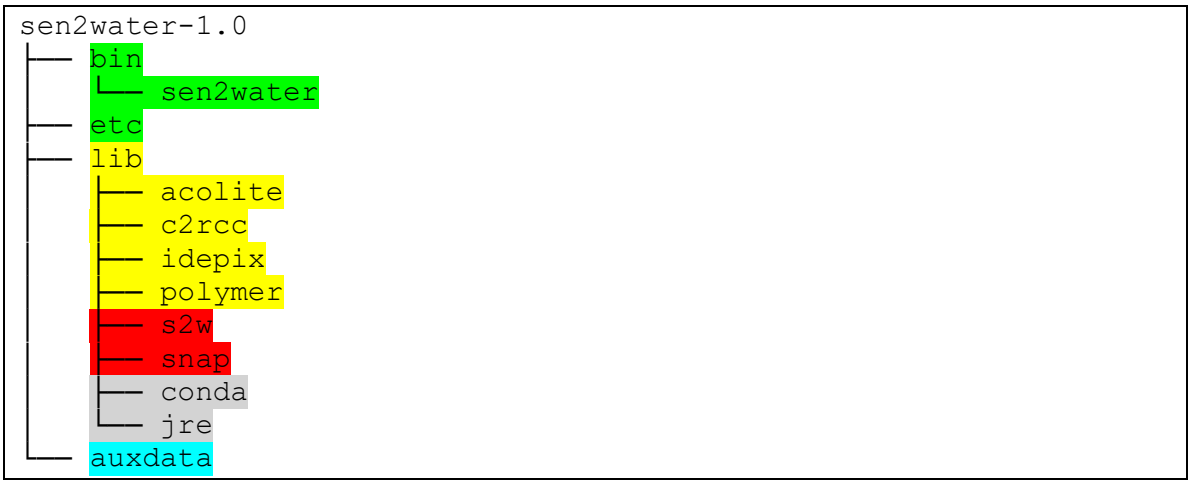

Figure 4-2: Software runtime structure of Sen2Water

- The lib subdirectories acolite, c2rcc, idepix, and polymer contain software packages of the four processors as available from their respective repositories.
- The lib subdirectories conda and jre contain the Python runtime and the Java runtime. The Python runtime is a miniconda environment that has been made relocatable with conda-pack.
- The lib subdirectory snap contains parts of the Sentinel Toolbox.
- The lib subdirectory s2w contains the implementation of blending inherited from HR-OC and of blending of Sen2Water as well as the NetCDF formatting code.
- The bin and etc directories contain the processor executable shell script, SNAP GPT graphs and processor parameter files.
- The auxdata directory contains static auxiliary data used by the processors.

This software installation directory can be deployed in an arbitrary system or user path, e.g. in /opt under Linux. Processing is started in a working directory:

```
cd /path/to/working/directory
```

Processing can be started with

```
/path/to/sen2water-1.0/bin/sen2water [<params>] <input> [<output>]
```

sen2water will write intermediate data and the output to the current working directory.

4.2.2 Sen2Water software installation package

The software installation package is a zip file that contains the software runtime structure starting with the directory sen2water-1.0 .

The installation package can be installed by

```
cd /path/to/installation/root
unzip /path/to/sen2water-linux-1.0.zip
```

The installation package is the delivery structure for the integration of Sen2Water into the combined Sentinel-2 Level 2 processor.

4.3 Combined Level 2 processor software structures

4.3.1 Merger

The merger will be provided as a small package that is based on a Sen2Cor installation. It uses the respective Python installation.

4.3.2 Workflow control script

The workflow control script will be provided as a small package that requires both a Sen2Cor installation and a Sen2Water installation.

Alternatively, the workflow control of Sen2Cor, Sen2Water, and the Merger are part of the PDGS integration that controls the call of Sen2Cor in the IPF today.

5 Sen2Water stand-alone

The functional elements and the structure of Sen2Water stand-alone is already described in section 3.2 and 4.2.

5.1 Target operating system dependent parts

The Java JRE and Python conda installations in the lib directory and the shell scripts are operating system dependent. There are packages for

- Windows
- Linux

5.2 Installation package

The Sen2Water software installation package is at the same time the Sen2Water stand-alone processor for Linux. It can be deployed and used independent of Sen2Cor.

A corresponding package is provided for Windows. The respective workflow control scripts are provided as Windows batch files.

The stand-alone version will have switched on functions to download auxiliary data, in particular DEM tiles required. This is assumed available in the local file system in the PDGS version. The download function can be disabled if the auxiliary data is available locally.

The stand-alone version development is subject to a later phase of the project.