Ocean Colour Multi-Mission Algorithm Prototype System (OMAPS)

# Input Output Data Description

**Ref:** D2.2
**Date:** 12/01/2022
**Issue:** 2.3
**For:** EUMETSAT

# Table of Contents

# Document Control

| Version No. | Release Date | Author/Contributor | Reason for issue |
|---|---|---|---|
| 0.1 | 08.11.19 | Danne, Brockmann, Mazeran, Krasemann, Jackson | initial version for review by partners |
| 0.2 | 02.12.19 | Danne, Brockmann, Mazeran, Krasemann, Jackson | revision after partial feedback from partners, release as **draft** to EUM for review |
| 0.3 | 06.03.20 | Danne, Brockmann, Mazeran, Krasemann, Jackson | revision of RR-AC, MATCHUPS, added SVC, added OC partially |
| 0.4 | 03.04.20 | Danne, Brockmann, Lebreton, Mazeran, Krasemann, Jackson | added VAL, completed OC part; intermediate release to EUM for review |
| 0.6 | 06.11.20 | Danne, Brockmann, Lebreton, Mazeran, Krasemann, Jackson, Dingle | baseline for a 'living document' which is always in line with the current software state |
| 1.0 | 08.01.21 | Danne, Brockmann, Lebreton, Mazeran, Krasemann, Jackson, Dingle | stay in line with the current software state |
| 1.1 | 31.07.21 | Danne, Brockmann, Lebreton, Mazeran, Krasemann, Jackson, Dingle | stay in line with the current software state |
| 2.0 | 15.10.21 | Danne, Jackson, Müller, Dingle, Mazeran, Brockmann, Lebreton, Krasemann | preliminary final version |
| 2.1 | 26.11.21 | Danne, Jackson, Müller, Dingle, Mazeran, Brockmann, Lebreton, Krasemann | revision after feedback from EUMETSAT |
| 2.2 | 7.12.21 | Danne, Jackson, Müller, Dingle, Mazeran, Brockmann, Lebreton, Krasemann | further revision after feedback from EUMETSAT |
| 2.2.1 | 12.01.22 | Danne, Jackson, Müller, Dingle, Mazeran, Brockmann, Lebreton, Krasemann | minor changes after feedback from EUMETSAT |
| 2.3 | 12.01.22 | Danne, Jackson, Müller, Dingle, Mazeran, Brockmann, Lebreton, Krasemann | merge of minor edits done by EUMETSAT |

# Applicable Documents

| ID | Document |
|---|---|
| AD-1 | EUMETSAT general SoW |
| AD-2 | EUMETSAT specific SoW for ocean colour prototype |
| AD-3 | ID 995318 Level 2 Product Monitoring –Evolution Studies Multi-mission Ocean Colour (OC) Prototyping algorithm – PML Proposal |
| AD-4 | OMAPS deliverable D1.1 Requirements Baseline Document for Offline Processor |
| AD-5 | OMAPS deliverable D1.2 Requirements Baseline Document for Ocean Colour Products |
| AD-6 | OMAPS deliverable D1.3 Source Code Installation and Software User Manual |
| AD-7 | OMAPS deliverable D2.1.3 ATBD for Atmospheric Correction Round Robin |
| AD-8 | OMAPS deliverable D2.3 Product Specification Document |
| AD-9 | OMAPS deliverable D4.1 Product Validation Plan |
| AD-10 | OMAPS deliverable D4.3 Product Validation and Evaluation Report |

# Reference Documents

| ID | Document |
|---|---|
| RD-1 | Franz, B. A., S. W. Bailey, P. J. Werdell, and C. R. McClain. 2007. "Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry." Appl Opt 46 (22): 5068-5082 |
| RD-2 | Bailey, Sean & Werdell, Jeremy. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. Remote Sensing of Environment. 102. 12-23. 10.1016/j.rse.2006.01.015. |
| RD-3 | Adam Lawson, Sherwin Ladner, Richard Crout, Christopher Wood, Robert Arnone, Jennifer Bowers, Paul Martinolich, and David Lewis "Establishing optimal matchup protocols between ocean color satellites and ground truth AeroNET-OC radiance", Proc. SPIE 11014, Ocean Sensing and Monitoring XI, 110140C (10 May 2019); https://doi.org/10.1117/12.2521062 |
| RD-4 | EUMETSAT 2021: Recommendations for Sentinel-3 OLCI Ocean Colour product validations in comparison with in situ measurements – Matchup Protocols. EUM/SEN3/DOC/19/1092968, v7, 17/05/2021 |
| RD-5 | Dagmar Müller, Hajo Krasemann, Robert J.W. Brewin, Carsten Brockmann, Pierre-Yves Deschamps, Roland Doerffer, Norman Fomferra, Bryan A. Franz, Mike G. Grant, Steve B. Groom, Frédéric Mélin, Trevor Platt, Peter Regner, Shubha Sathyendranath, François Steinmetz, John Swinton, The Ocean Colour Climate Change Initiative: II. Spatial and temporal homogeneity of satellite data retrieval due to systematic effects in atmospheric correction processors, Remote Sensing of Environment, Volume 162, 2015, Pages 257-270 |
| RD-6 | S3 Cal/Val plan, S3-PL-ESA-SY-0265, 2014 |
| RD-7 | OC_CCI - Phase Two, Product Validation and Algorithm Selection Report PVASR), |
| RD-8 | Dagmar Müller, Hajo Krasemann, Robert J.W. Brewin, Carsten Brockmann, Pierre-Yves Deschamps, Roland Doerffer, Norman Fomferra, Bryan A. Franz, Mike G. Grant, Steve B. Groom, Frédéric Mélin, Trevor Platt, Peter Regner, Shubha Sathyendranath, François Steinmetz, John Swinton. The Ocean Colour Climate Change Initiative: I. A methodology for assessing atmospheric correction processors based on in-situ measurements. Remote sens. Environ., Volume 162, 1 June 2015, Pages 242–256 |
| RD-9 | Sathyendranath, S, Brewin, RJW, Müller, D, Brockmann, C, Deschamps, P-Y, Doerffer, R, Fomferra, N, Franz, BA, Grant, MG, Hu C, Krasemann, H, Lee, Z, Maritorena, S, Devred, E, Mélin, F, Peters, M, Smyth, T, Steinmetz, F, Swinton, J, Werdell, J, Regner, P (2012) Ocean Colour Climate Change Initiative: Approach and Initial Results, IGARSS 2012: 2024-2027.http://dx.doi.org/10.1109/IGARSS.2012.6350979 |
| RD-10 | Jackson, T., R. Brewin, S. Sathyendranath, 2016: A round robin comparison of bio-optical algorithms applied to remotely-sensed reflectance data alongside fuzzy classification. Ocean Colour CCI PVASR Part2, available from https://esa-oceancolour-cci.org/sites/esa-oceancolour-cci.org/alfresco.php?file=103d0ddf-036d-420d-a8f9-2fc0cf54b159&name=OC-CCI-PVASR_PART2_In-Water_2014-01-22.pdf |
| RD-11 | Mazeran, C., Danne, O., Ruescas, A. (2020). Ocean Colour System Vicarious Calibration Tool: tool documentation (DOC-TOOL). Ref. EUM/19/SVCT/D2, version 1.0. Available at https://www.eumetsat.int/ocean-colour-system-vicarious-calibration-tool |
| RD-12 | EUMETSAT OCDB User Manual. Copyright 2019, Copernicus Revision 763d8907; https://ocdb.readthedocs.io/en/latest/index.html |
| RD-13 | Product Data Format Specification – OLCI Level 1 & Level 2 Instrument Products. ESA Ref: S3IPF.PDS.004, Issue 1.10, 28 May 2015. |
| RD-14 | Thuillier, G., M. Hers´e, P. C. Simon, D. Labs, H. Mandel, D. Gillotay, & T. Foujols. (2003). The solar spectral irradiance from 200 to 2400 nm as measured by the SOLSPEC spectrometer from the ATLAS 1-2-3 and EURECA missions. Solar Physics 214: 1-22. |
| RD-15 | Price-Whelan et al. (2018). The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. The Astronomical Journal 156: 123-141. |
| RD-16 | Steinmetz, F., and C. Mazeran: SACSO Atmospheric Aerosol Correction for Sentinel-3 Ocean Colour – Input Output Definition Document (IODD). Version 0.2, 19 Feb 2019. |

# List of Acronyms

| Acronym | Description |
|---|---|
| ADF | Auxiliary data file |
| ATBD | Algorithm Theoretical Baseline Document |
| BRDF | Bidirectional Reflectance Distribution Function |
| CBQ | Common Best Quality |
| CCI | Climate Change Initiative |
| C2RCC | Case 2 Regional Coastcolour Processor |
| EUM | Eumetsat |
| ESA | European Space Agency |
| FOV | Field of View |
| IBQ | Individual Best Quality |
| IODD | Input/Output Data Description |
| LUT | Look-up Table |
| MERIS | MEdium Resolution Imaging Spectrometer |
| OC-CCI | Climate Change Initiative Ocean Colour |
| OCDB | Copernicus Ocean Colour Reference database |
| OC-PROC | Ocean Colour Processor |
| OLCI | Ocean and Land Colour Instrument |
| OMAPS | Ocean Colour Multi-Mission Algorithm Prototype System |
| OWT | Optical Water Type |
| RR-AC | Round Robin Atmospheric Correction, a module of the *offline processor* subsystem of the Ocean Colour Multi-Mission Algorithm Prototype System |
| RR-IW | Round Robin In-Water, a module of the *offline processor* subsystem of the Ocean Colour Multi-Mission Algorithm Prototype System |
| SoW | Statement of Work |
| SVC | System Vicarious Calibration (general) or System Vicarious Calibration Module of the *offline processor* subsystem |
| TOA | Top of Atmosphere |
| VAL | Validation and Product Assessment Module of the *offline processor* subsystem |

## INDEX OF TABLES

## INDEX OF FIGURES

# 1 Purpose

This Input Output Data Description Document (IODD) aims at providing a complete list and a precise technical description of all products that will be used for / generated within the various project tasks arising from the project requirements [AD-4], [AD-8]. The document shall describe the format, structure and data ranges of all demanded input data, auxiliary files, generated output products and test and validation datasets.  It serves as a technical document that should clearly describe all the of datasets that are input into or output from the various modules contained with the online and offline processor modules.

This document is applicable to downstream deliverables.

# 2 Introduction

A prototype processing system is developed as the main outcome of this study contract. This prototype processing system is composed of two main sub-systems: a production chain, called the "*ocean colour processor*", the purpose of which is to systematically generate ocean colour products, and a so-called "*offline processor*" which performs supporting activities such as generation of coefficients required for processing or algorithm/product quality assessment. Both sub-systems will build upon the systems created under the Ocean-Colour Climate Change Initiative (OC-CCI) project. While the latter has been designed as a research platform, under this contract it will be further developed towards a prototype for an operational product-testing and operational processing system.

The main modules of both offline and ocean colour processor as well as the interaction of the two processor subsystems are illustrated in Figure 1.

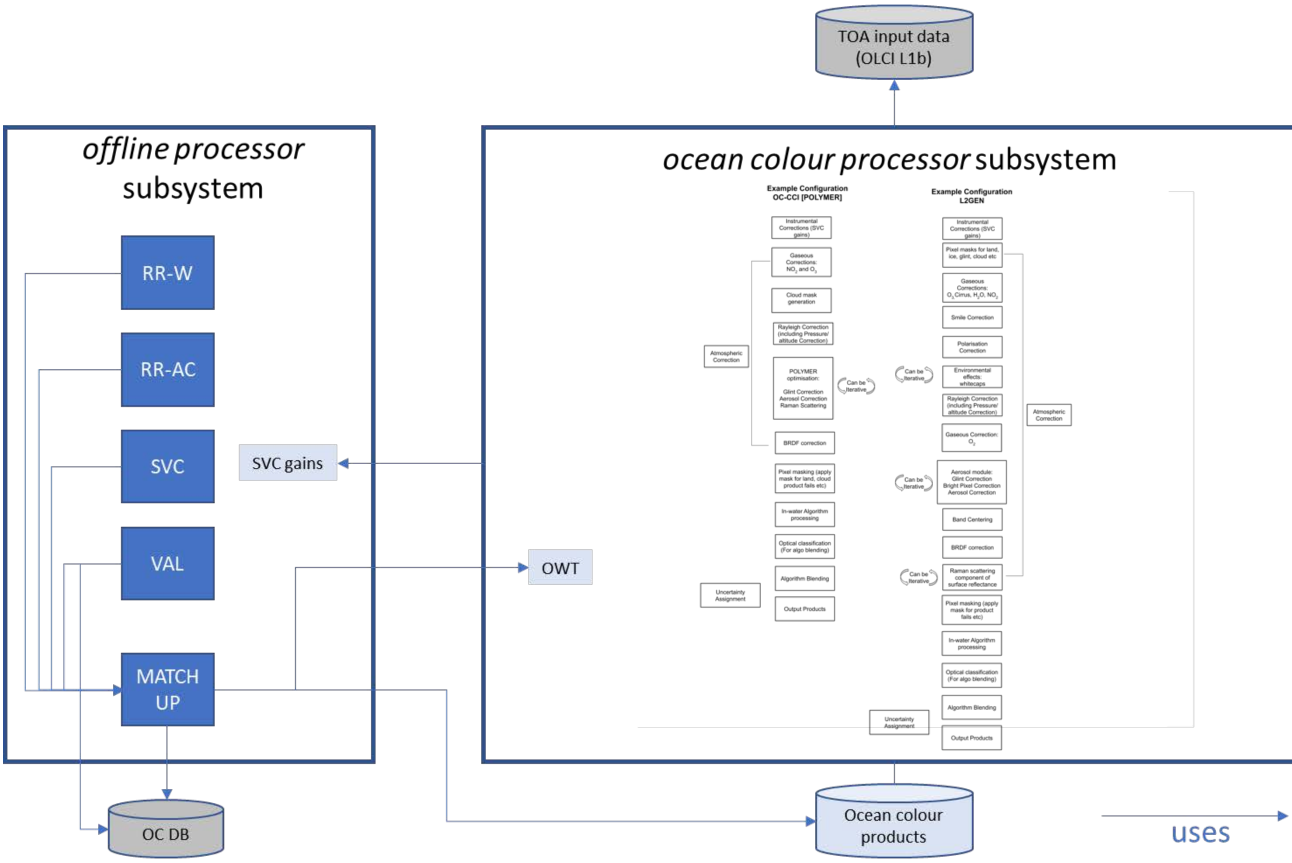



*Figure 1: Modules of the Ocean Colour Processor and the Offline Processor subsystem and its dependencies*

This figure reflects the main processor modules and tasks arising from the project requirements which are outlined in detail in [AD-4] and [AD-8]. In particular we have:

- Atmospheric Correction Round Robin (**RR-AC)**
- In-water Round Robin (**RR-IW)**
- System Vicarious Calibration (**SVC)**
- Validation and Quality Assessment (**VAL)**
- Match-up Generation **(MATCHUP)**
- Ocean Colour Processor **(OC-PROC)**

# 3 Input and Output Data Descriptions

The following subsections briefly describe each module within the subsystems, and list for each module the input data required by the module, and output data generated.

Input data comprise any data which is used as input by the module. These can be satellite data as the main data to be processed and which change frequently, as well as ancillary data such as thresholds, climatologies, gain coefficients, config files etc which may be totally static or change sometimes or change with every call of the module.

Output data comprise any data which is generated by the module and relevant for subsequent processing or usage by the user. This may include the main output such as surface reflectances or chlorophyll, but also logging and debugging information. Output data can be optional, e.g., additional marine variables as required by EUM SoW. In contrast to this, data which are temporarily written to disk and read by module are not included as they do not differ from data passed in memory and are deleted after usage.

## 3.1 Subsystem Ocean Colour Processor

### 3.1.1 Overview

The Ocean Colour Processor **(OC-PROC)** is fundamentally designed to systematically generate ocean colour products.

The prototype OC processor is modular in nature, allowing for modification or comparison of different methods or parameterizations at both the atmospheric correction and in-water product retrieval stages. A schematic of the proposed processing chain format is shown on in Figure 2. Given that different approaches to atmospheric correction (and different sensors) will require a different combination of modules it makes more sense to consider variant 'configuration' options rather than all possible module configurations. The validity of a given configuration will be testable using module tags. The tags are added sequentially for each module within the processor configuration and are not removed at any stage. This means that the order of operations and total list of operations can be checked for validity.

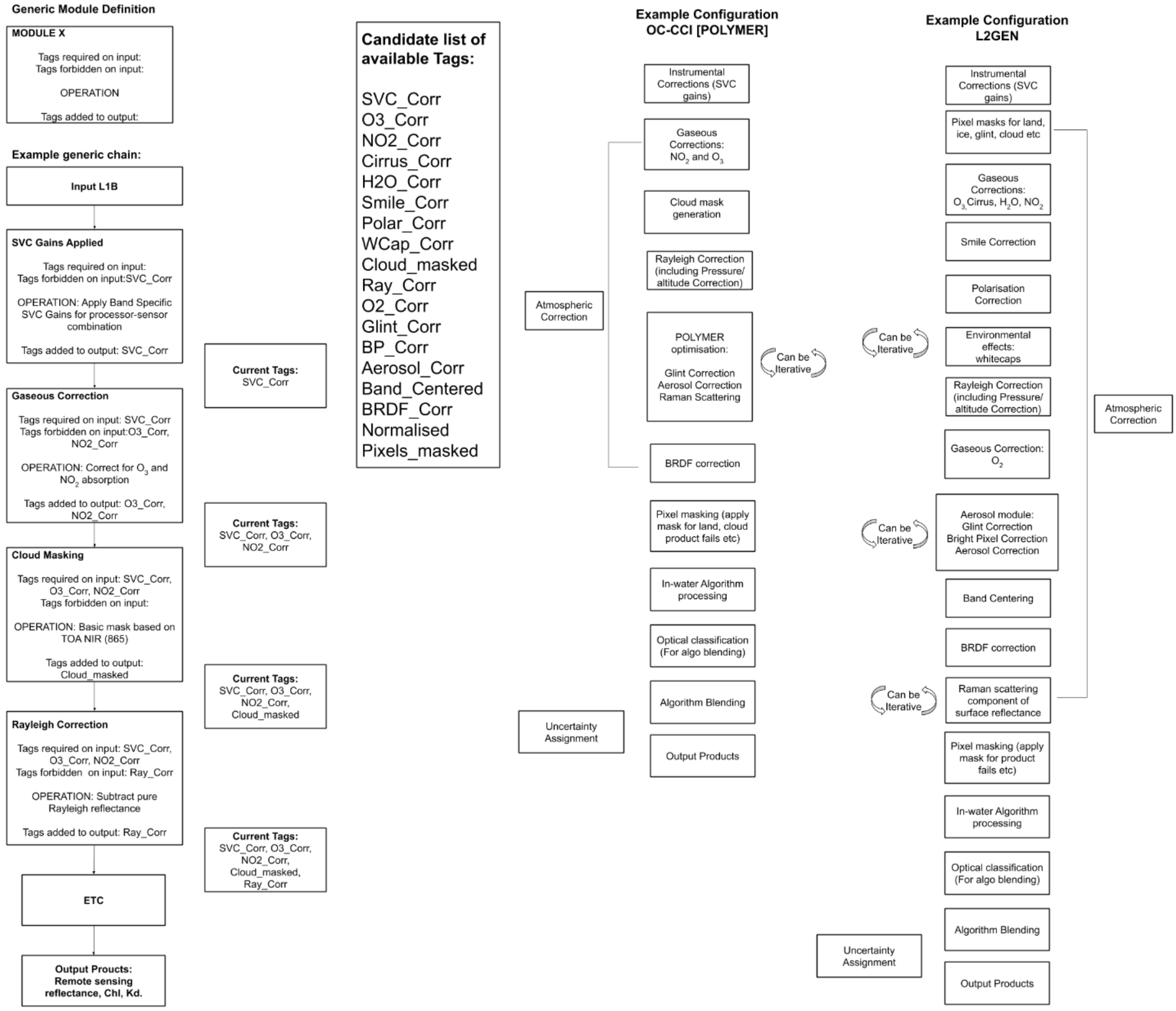



*Figure 2: Schematic diagram of modules within OC processing chain. Module format is shown on the left and potential prototype processor configurations are shown on the right. Taken from [AD-5].*

The overall structure of the ocean colour processing subsystem was already illustrated in Figure 1. The main submodules are in summary:

- Instrumental correction / SVC gains
- Smile correction
- Atmospheric correction, which can be subdivided into further submodules:
    - Gaseous correction
    - Polarisation Correction
    - Rayleigh correction
    - Environmental effects (such as whitecaps)
    - Cloud masking
    - Glint Correction
    - Bright Pixel Correction
    - Aerosol correction
    - Raman scattering and BRDF correction

- Pixel masking
- Optical classification
- In-water algorithms
- Product blending
- Product uncertainty
- File standardisation and checking

Below is a table of parameters used or generated within modules of the Ocean Colour Processor.

Table 3-1: Parameters relevant to the Ocean Colour Processor

| **Parameter name** | **Description** | **Dimensions** | **Unit** | **Type** |
|---|---|---|---|---|
| | ***Geometry*** | | | |
| sza | Solar zenith angle | row, column | degrees | float32 |
| vza | View zenith angle | row, column | degrees | float32 |
| raa | Relative azimuth angle (solar azimuth angle – view azimuth angle) (delta phi) | row, column | degrees | float32 |
| | ***Georeferencing*** | | | |
| Latitude | Latitude [-90, 90] | row, column | degrees | float32 |
| Longitude | Longitude [-180, 180] | row, column | degrees | float32 |
| | ***Radiometry*** | | | |
| Lt | Top of atmosphere radiance | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| wav | Effective measured wavelengths (per-pixel, per-band)<br>For sensors with fixed wavelength per band, it is equal to wav0. | band, row, column | nm | float32 |
| wav0 | Nominal wavelengths (per-band) | band | nm | float32 |
| Fsol | Solar irradiance | band, row, column or band | mW.m-2.nm-1 | float32 |
| rho_toa | Top of atmosphere reflectance<br>$\rho_{toa}(\lambda) = \frac{\pi \cdot L_{toa}(\lambda)}{\cos(\theta_s) \cdot F_0(\lambda)}$<br>Includes correction for sun-earth distance | band, row, column | dimensionless | float32 |
| | ***Ancillary data (may be embedded in level1 files or retrieved from datasets like ERA5 or SRTM3)*** | | | |
| altitude | Surface altitude (for processing of inland waters – used for calculating the Rayleigh optical depth)<br>Sources:<br>-Constant sea-level altitude<br>-SRTM3 | row, column | m | float32 |
| sea_level_pressure | Sea level pressure<br>Sources (first is default):<br>-OLCI Level1 (ECMWF forecast)<br>-NOAA NCEP Reanalysis<br>-ECMWF ERA-Interim Reanalysis<br>-ECMWF ERA5 Reanalysis | row, column | hPa | float32 |
| wind_speed | Wind speed module<br>Sources (first is default):<br>-OLCI Level1 (ECMWF forecast)<br>-NOAA NCEP Reanalysis<br>-ECMWF ERA-Interim Reanalysis<br>-ECMWF ERA5 Reanalysis | row, column | m/s | float32 |
| total_ozone | Total ozone column<br>Sources (first is default):<br>-OLCI Level1 (ECMWF forecast)<br>-TOMS/OMI<br>-ECMWF ERA-Interim Reanalysis<br>-ECMWF ERA5 Reanalysis | row, column | DU | float32 |

| | | | | |
|---|---|---|---|---|
| Total_NO2, Tropospheric_NO2 | Total and tropospheric NO2 columns<br>Source:<br>-NO2 OMI climatology 2004-2008<br>(will be replaced by 2014 climatology) | row, column | molecules/cm2 | float32 |
| total_CO2 | Total column CO2<br>Not yet used by Polymer | row, column | ppm | float32 |
| ***Intermediary products*** | | | | |
| rho_toa_svc | rho_toa, with SVC coefficients applied | band, row, column | dimensionless | float32 |
| rho_toa_gc | rho_toa_svc, corrected for gaseous absorption | band, row, column | dimensionless | float32 |
| rho_molgli | Pure Rayleigh reflectance, including the sun glint | band, row, column | dimensionless | float32 |
| rho_mol | Pure Rayleigh reflectance, without the sun glint (black surface) | band, row, column | dimensionless | float32 |
| Tmol | Pure Rayleigh transmittance | band, row, column | dimensionless | float32 |
| rho_prime | rho_toa_gc- rho_molgli | band, row, column | dimensionless | float32 |
| rho_prime_noglint | rho_toa_gc- rho_mol | band, row, column | dimensionless | float32 |
| rhowmod | Water reflectance: output of the model with the final values | band, row, column | dimensionless | float32 |
| rhow | Water reflectance, without spectral and directional normalization | band, row, column | dimensionless | float32 |
| *Es* | Solar Irradiance at Surface | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| Lt_svc | TOA radiance after SVC correction | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| Lt_gc | TOA radiance after gaseous correction | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| Lt_sc | TOA radiance after smile correction. | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| Lt_pc | TOA Radiance after polarisation correction | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| Lt_f | TOA radiance from Whitecaps | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| Lt_fc | TOA after correction for whitecaps | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| Lr | Rayleigh Radiance | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| Lt_rc | TOA radiance after Rayleigh Radiance Correction | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| Lt_gli | TOA glint radiance (from Cox and Munk) | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| Lt_glc | TOA radiance after glint correction | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| Lt_bpc | TOA radiance after NIR bright pixel correction | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| *Lt_w* | TOA detected Water Leaving Radiance | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| *Lw* | Water Leaving Radiance at surface | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| *Tau_r* | *Rayleigh optical thicknesses* | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| *Tau_a* | *Retrieved aerosol optical thicknesses* | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| *tg_sen* | *Gaseous Transmittance, Surface to Sensor* | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| *tg_sol* | *Gaseous Transmittance, Sun to Surface* | band, row, column | mW.m-2.sr-1.nm-1 | float32 |
| ***Output vector of the minimization (parameters of the water reflectance model)*** | | | | |
| logchl | 1st component of the solution vector:<br>log10 of the chlorophyll concentration | row, column | log10(mg.m-3) | float32 |
| logfb | 2$^{nd}$ component of the solution vector:<br>log10 of the backscattering ratio | row, column | dimensionless | float32 |
| ***Final products (level2)*** | | | | |
| rhown | Water reflectance, spectrally and directionally normalized. Sometimes improperly referred to as water “leaving” reflectance. | band, row, column | dimensionless | float32 |
| Rrs | Remote sensing reflectance, spectrally and directionally normalized. | band, row column | sr-1 | float32 |
| ***Flags*** | | | | |
| flags | Binary flags, which serve either as input or as output. | row, column | Not applicable | uint32 |

The following sections cover the inputs and outputs relevant to each component module available for a processor configuration.

### 3.1.2 Atmospheric Correction

Currently the processor does not have a fully modular breakdown of the atmospheric correction. Though the steps covered in Figure 2 are undertaken by atmospheric correction schemes such as POLYMER and L2gen, the system cannot currently break those steps apart. As such, the atmospheric correction step should currently be considered as 1 'module' with inputs and outputs.

#### 3.1.2.1 Input Data

The function of this module is to provide surface reflectance values by applying the following steps, as required, for atmospheric correction. The input data required is listed below, separated by the operation it is used for:

*Apply system vicarious calibration:*

- Input L1B TOA radiance data
- SVC gains LUT for the given atmospheric processor configuration

*Apply gaseous correction:*

- TOA radiances (for example $O_2$ absorption lines are heterogeneously distributed across band O16)
- relevant LUT for gaseous properties
- estimate of gaseous concentrations
- viewing and solar geometry
- estimate of aerosol thickness

*Apply smile correction:*

- Knowledge of in-FOV variation of wavelength for OLCI cameras
- Solar flux
- Date (for solar flux correction)

*Apply polarisation correction:*

- LUT file for Sensor sensitivity to polarisation

*Apply basic cloud masking:*

- Method for cloud top height estimation
- Viewing and solar geometry information

*Apply Rayleigh correction:*

- Rayleigh scattering LUTs
- Estimates of surface altitude and atmospheric pressure

*Apply whitecap correction:*

- Windspeed data
- Atmospheric Pressure
- Whitecap spectral shape

*Apply glint correction:*

- Glitter radiance
- air mass

*Apply aerosol correction:*

- Viewing and solar geometry information
- LUTs of aerosol properties

*Apply raman scattering and brdf correction:*

- Raman excitation response functions per sensor band
- LUTs for Raman scattering properties
- Downwelling irradiance data

*Table 3-2: Main properties of OC-Atmospheric correction input/output data set*

| ***Processing step*** | **Atmospheric correction** |
|---|---|
| ***Tags required on input data*** | |
| ***Tags forbidden on input data*** | SVC_Corr |
| ***Additional files / configuration*** | User configurable:<br>• Sensor-processor specific gains per band of sensor - provided via 'svc_gains' config item in the configuration file<br>Provided by AC processor:<br>• Ancillary information on gaseous concentrations<br>• LUT for sensor band characteristics and detector configuration<br>• LUT for polarisation sensitivity<br>• Cloud identification criteria or identification scheme<br>• Rayleigh scattering LUT and pressure data<br>• Ancillary environmental data (windspeed, pressure)<br>• Aerosol properties LUT<br>• Optical properties of pure water<br>• Raman response LUT<br>• Solar spectrum<br>• Solar flux<br>• BRDF LUT (f/q table) |
| ***Data File Format*** | Netcdf (TOA radiance, Ancillary environmental data),<br>hdf (),<br>text files (csv or tab delimited) [LUT for sensor band characteristics, polarisation sensitivity, BRDF, optical properties of pure water, Raman response) |
| ***Input variable(s)*** | Lt (TOA radiance) |
| ***Output variable(s)*** | nLw (Normalised water leaving radiance)<br>Rrs (Remote sensing reflectance) |
| ***Dimensions*** | Same as input file |
| ***Size per file*** | Approximately same as input file |
| ***Amount of files*** | NA |
| ***Tags added to output data*** | SVC_Corr, Ozone_Corr, NO2_Corr, Ray_Corr, WCap_Corr, Glint_Corr, Aerosol_Corr |

### 3.1.3 Pixel masking

The function of this module is to remove pixels classified as land or flagged by preceding modules.

#### *3.1.3.1 Input Data*

- Product Flags
- Configuration file of which flags to mask

*Table 3-3: Main properties of OC-PROC Filtering and masking input/output data set*

| **Processing step** | **Filtering and masking** |
|---|---|
| **Tags required on input data** | SVC_Corr, Ozone_Corr, NO2_Corr, Ray_Corr, WCap_Corr, Glint_Corr, Aerosol_Corr |
| **Tags forbidden on input data** | Pixels_masked |
| **Additional files/configuration** | Mask configuration file<br>Flag masks (if additional flags are to be applied) |
| **Data File Format** | Netcdf |
| **Input variable(s)** | Rrs (Remote sensing reflectance) [these will be masked]<br>mask flags |
| **Output variable(s)** | Rrs (Remote sensing reflectance) |
| **Dimensions** | Same as input file |
| **Size per file** | Approximately same as input file |
| **Amount of files** | NA |
| **Tags added to output data** | Pixels_masked |

### 3.1.4 Optical classification

The function of this module is to classify the surface reflectance spectra against a defined optical water class set. This will allow optimal algorithm blending and uncertainty estimation.

#### 3.1.4.1 Input Data

- LUTs for optical water class set
- Level-2, remotely sensed reflectance data

*Table 3-4: Main properties of OC-PROC Optical classification input/output data set*

| **Processing step** | **Optical classification** |
|---|---|
| **Tags required on input data** | SVC_Corr, Ozone_Corr, NO2_Corr, Ray_Corr, WCap_Corr, Glint_Corr, Aerosol_Corr, BRDF_Corr |
| **Tags forbidden on input data** | |
| **Additional files/configuration** | Optical water class set LUT |
| **Data File Format** | Netcdf (both Rrs and LUT) |
| **Input variable(s)** | Rrs |
| **Output variable(s)** | Optical water class memberships<br>OWT_flag |
| **Dimensions** | Same as input file |
| **Size per file** | Approximately same as input file |
| **Amount of files** | NA |
| **Tags added to output data** | |

### 3.1.5 In-water algorithms

The function of this module is to derive in-water products (Chlorophyll-a and Kd490) from Rrs data.

#### *3.1.5.1 Input Data*

- Level-2, remotely sensed, surface water reflectance data
- Configuration file of which flags to mask

*Table 3-5: Main properties of OC-PROC In-water algorithms input/output data set*

| | |
|---|---|
| ***Processing step*** | **In water products** |
| ***Tags required on input data*** | SVC_Corr, Ozone_Corr, NO2_Corr, Ray_Corr, WCap_Corr, Glint_Corr, Aerosol_Corr, BRDF_Corr |
| ***Tags forbidden on input data*** | |
| ***Additional files/configuration*** | |
| ***Data File Format*** | Netcdf |
| ***Input variable(s)*** | Rrs |
| ***Output variable(s)*** | Chl_a products (multiple candidate algorithms)<br>Kd_490<br>Flags (product failure) |
| ***Dimensions*** | Same as input file |
| ***Size per file*** | Approximately same as input file |
| ***Amount of files*** | NA |
| ***Tags added to output data*** | In_water_products |

### 3.1.6 Product blending

The function of this module is to provide a seamlessly blended product that makes use of the best candidate algorithm for the optical conditions on a pixel-by-pixel basis.

#### *3.1.6.1 Input Data*

- Configuration file of optimal algorithm per optical class
- Configuration file of which flags to mask

*Table 3-6: Main properties of OC-PROC Product blending input data set*

| ***Processing step*** | **Product blending** |
|---|---|
| ***Tags required on input data*** | SVC_Corr, Ozone_Corr, NO2_Corr, Ray_Corr, WCap_Corr, Glint_Corr, Aerosol_Corr, BRDF_Corr |
| ***Tags forbidden on input data*** | |
| ***Additional files/configuration*** | LUT of blending scheme for algorithm weighting |
| ***Data File Format*** | Netcdf (data file)<br>Text file (blending scheme weights) |
| ***Input variable(s)*** | Rrs<br>OWT memberships<br>Chl_a products (multiple candidate algorithms) |
| ***Output variable(s)*** | Chl_a blended product |
| ***Dimensions*** | Same as input file |
| ***Size per file*** | Approximately same as input file |
| ***Amount of files*** | NA |
| ***Tags added to output data*** | Blended_products |

### 3.1.7 Product uncertainty

The function of this module is to provide information on the potential bias and accuracy of the L2 products.

#### 3.1.7.1 Input Data

- Optical water class memberships
- LUTs for per wc performance for given products

*Table 3-7: Main properties of OC-PROC Product uncertainty input data set <…>*

| ***Processing step*** | **Product uncertainty** |
|---|---|
| ***Tags required on input data*** | SVC_Corr, Ozone_Corr, NO2_Corr, Ray_Corr, WCap_Corr, Glint_Corr, Aerosol_Corr, BRDF_Corr |
| ***Tags forbidden on input data*** | |
| ***Additional files/configuration*** | LUT of per OWT performance characteristics |
| ***Data File Format*** | Netcdf |
| ***Input variable(s)*** | Rrs<br>OWT |
| ***Output variable(s)*** | Product uncertainty estimates (bias and accuracy) |
| ***Dimensions*** | Same as input file |
| ***Size per file*** | Approximately same as input file |
| ***Amount of files*** | NA |
| ***Tags added to output data*** | Uncertainty_added |

### 3.1.8 File standardisation and checking

The function of this module is to last file additions or updates relating to meta-data, versioning, standardisation, and product set merging.

#### *3.1.8.1 Input Data*

- Level-2 product files
- Configuration file for meta-data requirements

*Table 3-8: Main properties of OC-PROC File standardisation and checking input data set <…>*

| ***Processing step*** | **File standardisation and checking** |
|---|---|
| ***Tags required on input data*** | SVC_Corr, Ozone_Corr, NO2_Corr, Ray_Corr, WCap_Corr, Glint_Corr, Aerosol_Corr, BRDF_Corr, Rrs |
| ***Tags forbidden on input data*** | Standardised |
| ***Additional files/configuration*** | Metadata configuration file |
| ***Data File Format*** | Netcdf |
| ***Input variable(s)*** | |
| ***Output variable(s)*** | |
| ***Dimensions*** | Same as input file |
| ***Size per file*** | Approximately same as input file |
| ***Amount of files*** | NA |
| ***Tags added to output data*** | Standardised |

### 3.1.9 Online Processor Output

The resulting file after all stages of the Online Processor have completed will serve as the Ocean Colour input to the Offline Processor.

#### *3.1.9.1 Output Data*

*Table 3-9: Main properties of OC output data file*

| | |
|---|---|
| ***Tags required on input data*** | SVC_Corr, Ozone_Corr, NO2_Corr, Ray_Corr, WCap_Corr, Glint_Corr, Aerosol_Corr, BRDF_Corr, Rrs, Standardised |
| ***Additional files/configuration*** | N/A, file is self-describing |
| ***Data File Format*** | Netcdf |
| ***Output variable(s)*** | Rrs<br>Water reflectance<br>OWT memberships<br>Chl_a products (multiple candidate algorithms)<br>Blended Chl product |
| ***Dimensions*** | Same as input file |
| ***Size per file*** | variable |
| ***Amount of files*** | N/A |

*Table 3-10: Output variable description from Online Processor*

| Parameter name | Description | Dimensions | Unit | Type |
|---|---|---|---|---|
| | | ***Georeferencing*** | | |
| Latitude | Latitude [-90, 90] | row, column | degrees | float32 |
| Longitude | Longitude [-180, 180] | row, column | degrees | float32 |
| | | ***Output Products*** | | |
| Rw_<band> | Water reflectance | row column | Dimensionless, fully normalised | float32 |
| Rrs_<band> | Remote sensing reflectance, spectrally and directionally normalized. | row column | sr-1 | float32 |
| water_class<class> | Water class membership | row, column | N/A | float32 |
| chlor_<algorithm> | Chlor_a products from multiple candidate algorithms (currently: ocx, oc2, oc3, oc4, oci, oci2. huCI, huCI2) | row, column | $mg/m^{-3}$ | float32 |
| chlor_a_blended | Chlor_a product comprising candidate algorithms blended by waterclass membership | row, column | $mg/m^{-3}$ | float32 |

## 3.2 Subsystem Offline Processor

### 3.2.1 Module Match-up Generation

#### *3.2.1.1 Overview*

The basic concept of match-up generation is the following: two data sets exist for reference ground data and ocean colour satellite data. The match-up dataset is a new set conformed of ground-satellite measurement pairs which can be considered as “simultaneous” (in space and time) and representative of the same quantity or population In practice this can be very difficult because the measurement methods are different, the spatial extent of measurements are different, and perfect co-incidence is practically never given. Environmental aspects such as spatial heterogeneity of the water or atmosphere, or contamination by clouds, are further complications. Best practices exist to optimise the match-up procedure. Key reference publication as [RD-1] and [RD-2] which contains a detailed description of the in-situ match-up process. An update has been recently published [RD-3]. Most relevant for this study is [RD-4], as it has been developed specifically for OLCI on Sentinel 3 under the coordination of EUMETSAT.

The Match-up Generation Module (**MATCHUP**) for OMAPS basically follows the methodology developed by [RD-2], i.e., it contains the following steps to generate a matchup:

1. Select a valid reference in situ measurement; this has been measured at time *tref* and at location ($lat_{ref}$, $lon_{ref}$).
2. Find corresponding satellite product(s) according to time and location criteria.
3. Extract $N_{macro}$ x $N_{macro}$ macro pixel window around the location ($lat_{ref}$, $lon_{ref}$).
4. Perform filtering of valid pixels within the macro pixel; exclude matchups if macro-pixel quality criteria are not met.
5. Calculate statistical quantities (in particular, a measure of central value and a measure of type A uncertainty – e.g. standard deviation) of the satellite product for the macro pixel window.
6. If applicable, apply additional filter criteria.
7. Provide a match-up record composed of pairs of ground-reference and satellite measurements, (including all the relevant statistical information extracted from the window).

This procedure can be adapted to specific needs by user defined parameters in order to fulfil the requirements given in [AD-4]. Defaults for the relevant parameters are set as defined in [RD-4]. The architecture of the OMAPS MATCHUP Module is similar to EUMETSAT's MDB generation scheme as described in [RD-12] but has been adapted to the specific needs of OMAPS. E.g., the OMAPS MATCHUP Module allows for the generation of matchups using L2 ocean colour satellite products generated by the OMAPS OC Processor (section 3.1, probably the most common use case). In the same way as the MDB workflow, for the OMAPS

MATCHUP workflow it  is also possible to use L1 TOA radiance products to generate L1 matchups as needed by the SVC Module (section 3.2.4). The high-level workflow of the module is illustrated in Figure 3.

The required input for the matchup extraction is:

- Satellite data sources:
  - OLCI Level-1 TOA radiance products for the generation of L1 matchups
  - OLCI Level-2 Water products generated by the OMAPS Ocean Colour processor, using different Atmospheric Correction candidates. These products are described in detail in section 3.1.9.
- In-situ data source: The Copernicus OCDB [RD-12]
- Configuration parameters / user options: Provided as configuration file (ini format) generated by the Configuration Tool, described in more detail in the Software User Manual [AD-6].

Intermediate products generated from these input data sources are:

- OLCI minifiles on $N_{macro}xN_{macro}$ macro pixel window around the selected site, includes original data of the matchup variables specified
- In situ ‘extraction files’ (CSV format), containing the results of an OCDB query for given reference time, location and variables of interest

The final Matchup Database (MDB) products are netCDF files, CSV summary files, or both (depending on user option), including the matchups between OLCI and the in-situ data.

Format and content of the products mentioned above as well as the main preparation steps are given in more detail below.

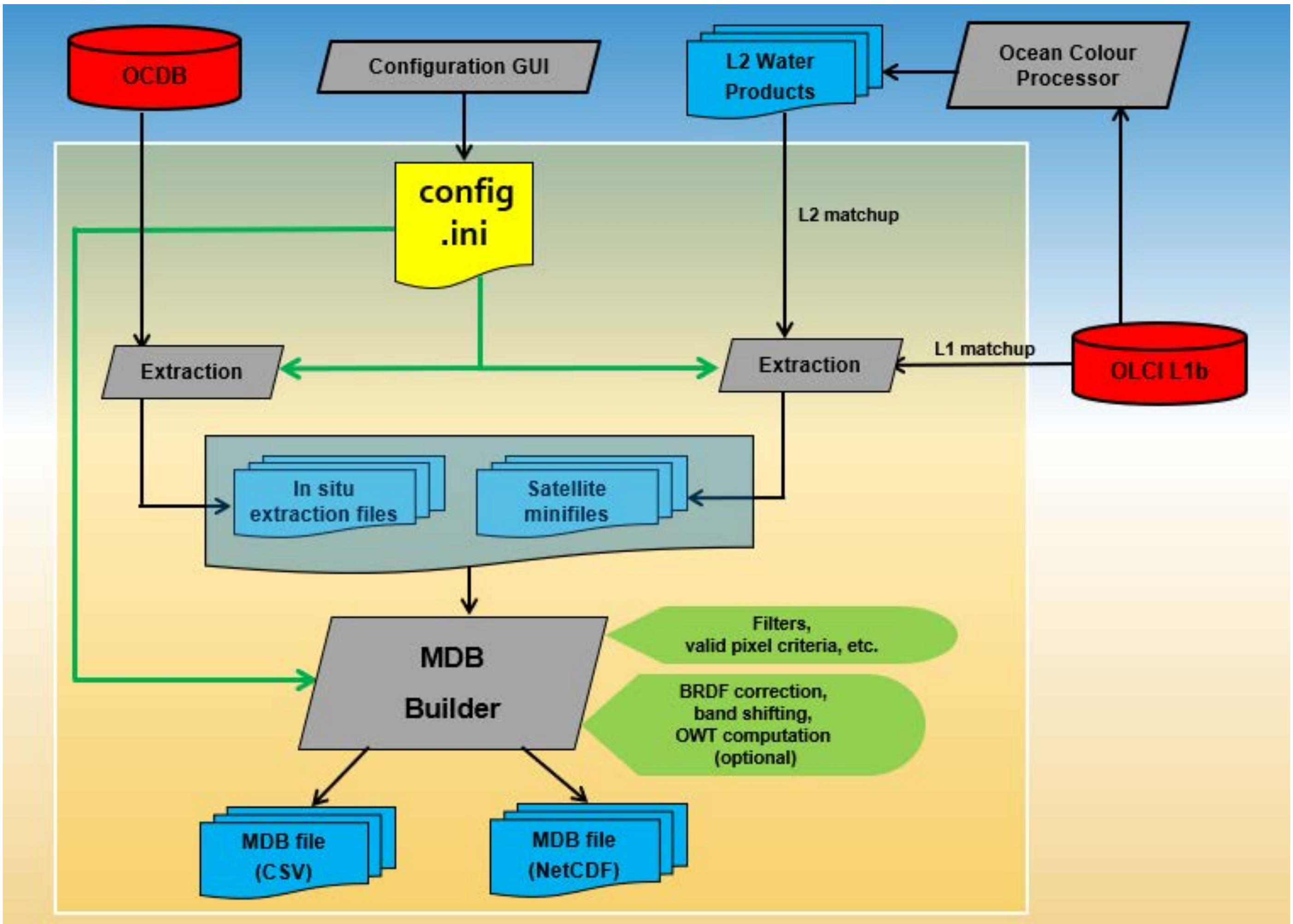



*Figure 3: High level workflow of the MATCHUP Module.*

### *3.2.1.2 Input Data*

#### 3.2.1.2.1 OLCI Level-1 TOA radiance products

The format and content of OLCI Level-1 TOA radiance products is described in full detail in [RD-13]. For OMAPS purposes, these products are accessed from EUMETSAT's CODA repository, https://coda.eumetsat.int. Details can be found in the corresponding User Manual, https://coda.eumetsat.int/manual/CODA-user-manual.pdf.

#### 3.2.1.2.2 OLCI Level-2 water products

The format and content of OLCI Level-2 water products as generated by the OMAPS Ocean Colour processor has been described in detail in section 3.1.9.1, see tables from Table 3-2 to Table 3-10

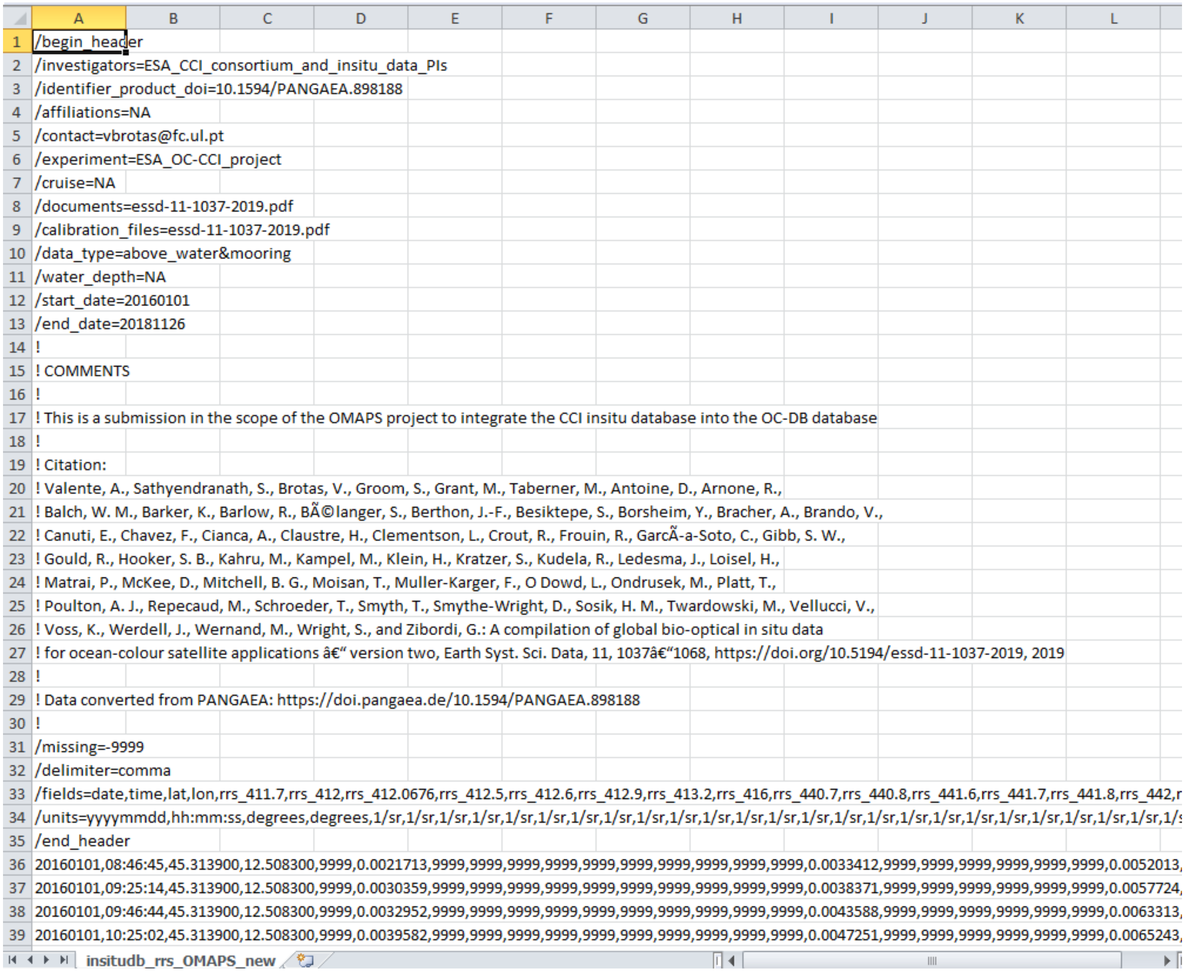
#### 3.2.1.2.3 In-situ data: Copernicus OCDB

The in situ data gathered from OCDB follows protocols and quality standards as described in detail in section 'Preparing in situ data: other sources' in https://ocdb.readthedocs.io/en/latest/ocdb-MDB-user-manual.html#mdb-files-content. This means that the particular content and available variables can be different for the given datasets. Figure 4 and Figure 5 show two examples of datasets taken from the OCDB. Both datasets originate from the OC-CCI project. The first dataset contains chl_a and related variables, the second dataset contains Rrs at various wavelengths.

```
/begin_header
/investigators=ESA_CCI_consortium_and_insitu_data_PIs
/identifier_product_doi=10.1594/PANGAEA.898188
/affiliations=NA
/contact=vbrotas@fc.ul.pt
/experiment=ESA_OC-CCI_project
/cruise=NA
/documents=essd-11-1037-2019.pdf
/calibration_files=essd-11-1037-2019.pdf
/data_type=above_water&mooring
/water_depth=NA
/start_date=20160107
/end_date=20170120
!
! COMMENTS
!
! This is a submission in the scope of the OMAPS project to integrate the CCI insitu database into the OC-DB database
!
! Citation:
! Valente, A., Sathyendranath, S., Brotas, V., Groom, S., Grant, M., Taberner, M., Antoine, D., Arnone, R.,
! Balch, W. M., Barker, K., Barlow, R., BÃ©langer, S., Berthon, J.-F., Besiktepe, S., Borsheim, Y., Bracher, A., Brando, V.,
! Canuti, E., Chavez, F., Cianca, A., Claustre, H., Clementson, L., Crout, R., Frouin, R., GarcÃ-a-Soto, C., Gibb, S. W.,
! Gould, R., Hooker, S. B., Kahru, M., Kampel, M., Klein, H., Kratzer, S., Kudela, R., Ledesma, J., Loisel, H.,
! Matrai, P., McKee, D., Mitchell, B. G., Moisan, T., Muller-Karger, F., O Dowd, L., Ondrusek, M., Platt, T.,
! Poulton, A. J., Repecaud, M., Schroeder, T., Smyth, T., Smythe-Wright, D., Sosik, H. M., Twardowski, M., Vellucci, V.,
! Voss, K., Werdell, J., Wernand, M., Wright, S., and Zibordi, G.: A compilation of global bio-optical in situ data
! for ocean-colour satellite applications â€“ version two, Earth Syst. Sci. Data, 11, 1037â€“1068, https://doi.org/10.5194/essd-11-1037-2019, 2019
!
! Data converted from PANGAEA: https://doi.pangaea.de/10.1594/PANGAEA.898188
!
/missing=-9999
/delimiter=comma
/fields=date,time,lat,lon,chla_hplc,chla_fluor,chla_hplc_dataset,chla_hplc_subdataset,chla_hplc_contributor,chla_fluor_dataset,chla_fluor_subdataset,chla_fluor_contributor
/units=yyyymmdd,hh:mm:ss,degrees,degrees,mg/m^3,mg/m^3,none,none,none,none,none,none
/end_header
20160107,18:18:42,32.953330,-117.306670,-9999,0.58533,-9999,-9999,-9999,calcofi,calcofi_#09330267_#34145,Ralf_Goericke
20160107,21:02:01,32.952000,-117.285670,-9999,0.84875,-9999,-9999,-9999,calcofi,calcofi_#09340264_#34160,Ralf_Goericke
20160107,23:49:31,33.238170,-117.464000,-9999,0.62,-9999,-9999,-9999,calcofi,calcofi_#09170264_#34144,Ralf_Goericke
20160108,02:52:03,32.913830,-117.390330,-9999,0.50733,-9999,-9999,-9999,calcofi,calcofi_#09330280_#34146,Ralf_Goericke
```

insitudb_chla_OMAPS_new

*Figure 4: In situ dataset in the OCDB, containing chl_a and related variables. The dataset originates from the OC-CCI project.*



```
/begin_header
/investigators=ESA_CCI_consortium_and_insitu_data_PIs
/identifier_product_doi=10.1594/PANGAEA.898188
/affiliations=NA
/contact=vbrotas@fc.ul.pt
/experiment=ESA_OC-CCI_project
/cruise=NA
/documents=essd-11-1037-2019.pdf
/calibration_files=essd-11-1037-2019.pdf
/data_type=above_water&mooring
/water_depth=NA
/start_date=20160101
/end_date=20181126
!
! COMMENTS
!
! This is a submission in the scope of the OMAPS project to integrate the CCI insitu database into the OC-DB database
!
! Citation:
! Valente, A., Sathyendranath, S., Brotas, V., Groom, S., Grant, M., Taberner, M., Antoine, D., Arnone, R.,
! Balch, W. M., Barker, K., Barlow, R., BÃ©langer, S., Berthon, J.-F., Besiktepe, S., Borsheim, Y., Bracher, A., Brando, V.,
! Canuti, E., Chavez, F., Cianca, A., Claustre, H., Clementson, L., Crout, R., Frouin, R., GarcÃ-a-Soto, C., Gibb, S. W.,
! Gould, R., Hooker, S. B., Kahru, M., Kampel, M., Klein, H., Kratzer, S., Kudela, R., Ledesma, J., Loisel, H.,
! Matrai, P., McKee, D., Mitchell, B. G., Moisan, T., Muller-Karger, F., O Dowd, L., Ondrusek, M., Platt, T.,
! Poulton, A. J., Repecaud, M., Schroeder, T., Smyth, T., Smythe-Wright, D., Sosik, H. M., Twardowski, M., Vellucci, V.,
! Voss, K., Werdell, J., Wernand, M., Wright, S., and Zibordi, G.: A compilation of global bio-optical in situ data
! for ocean-colour satellite applications â€“ version two, Earth Syst. Sci. Data, 11, 1037â€“1068, https://doi.org/10.5194/essd-11-1037-2019, 2019
!
! Data converted from PANGAEA: https://doi.pangaea.de/10.1594/PANGAEA.898188
!
/missing=-9999
/delimiter=comma
/fields=date,time,lat,lon,rrs_411.7,rrs_412,rrs_412.0676,rrs_412.5,rrs_412.6,rrs_412.9,rrs_413.2,rrs_416,rrs_440.7,rrs_440.8,rrs_441.6,rrs_441.7,rrs_441.8,rrs_442,r
/units=yyyymmdd,hh:mm:ss,degrees,degrees,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/s
/end_header
20160101,08:46:45,45.313900,12.508300,9999,0.0021713,9999,9999,9999,9999,9999,9999,9999,9999,9999,9999,0.0033412,9999,9999,9999,9999,9999,9999,0.0052013,
20160101,09:25:14,45.313900,12.508300,9999,0.0030359,9999,9999,9999,9999,9999,9999,9999,9999,9999,9999,0.0038371,9999,9999,9999,9999,9999,9999,0.0057724,
20160101,09:46:44,45.313900,12.508300,9999,0.0032952,9999,9999,9999,9999,9999,9999,9999,9999,9999,9999,0.0043588,9999,9999,9999,9999,9999,9999,0.0063313,
20160101,10:25:02,45.313900,12.508300,9999,0.0039582,9999,9999,9999,9999,9999,9999,9999,9999,9999,9999,0.0047251,9999,9999,9999,9999,9999,9999,0.0065243,
```



*Figure 5: In situ dataset in the OCDB, containing Rrs at various wavelengths. The dataset originates from the OC-CCI project.*

### 3.2.1.2.4 Configuration file

The configuration parameters / user options for the matchup generation are provided as configuration file (ini format). This configuration file is most conveniently generated by the Configuration GUI which is described in more detail in the SUM [AD-5]. Table 3-11 lists the available user options and their meaning, Table 3-12 lists the default valid expressions for the supported L2 processors, and Figure 6 shows an example for a configuration file used with a matchup test generation.

*Table 3-11: Configuration parameters used in the OMAPS matchup generation*

| PARAMETER | DESCRIPTION | POSSIBLE VALUES | DEFAULT |
|---|---|---|---|
| ***satellite_input_dir*** | Full path to satellite input data root directory | /path/to/directory | |
| ***sensors*** | Sensor (currently only OLCI) | OLCI | OLCI |
| ***platforms*** | Platforms (S3 A or B or both) | A, B, A and B | A and B |
| ***processing_level*** | Processing level (1 or 2) | 1, 2 | 1 |
| ***l2_processor*** | L2 processor name (only relevant for processing level 2) | POLYMER, SACSO, IPF (,L2GEN) | POLYMER |
| ***timeliness*** | Set NT for non-time critical products, NR for real-time products | NR, NT | NT |
| ***resolution*** | Satellite data resolution (OLCI Full or Reduced) | F, R | F |
| ***extraction_dir*** | Full path to OCDB and minifiles extraction root directory | /path/to/directory | |
| ***mdb_output_dir*** | Full path to MDB output root directory | /path/to/directory | |
| ***matchup_write_netcdf*** | If set to True, MDB netCDF product is generated | True, False | True |
| ***matchup_write_csv*** | If set to True, MDB CSV summary file is generated | True, False | True |
| ***matchup_write_owt*** | If set to True, OWT variables are written to MDB (if OWT normalisation is applied) | True, False | True |
| ***owt_normalisation*** | If set to True, OWT normalisation is applied | True, False | True |
| ***log_dir*** | Location of MDB generation log file | /path/to/directory | |
| ***processing_label*** | Label for this production (used as subdirectory name in extraction path, must not be empty | String | MDB_GENERATION |
| ***ocdb_username*** | Username of OCDB account | String | |
| ***ocdb_password*** | Password of OCDB account | String | |
| ***matchup_variables*** | Semicolon separated list of parameters match sets for which the matchups shall be extracted. For each match set, in-situ variable(s) are in a comma separated list to the left of a colon and satellite variable(s) go in a comma separated list on the right.<br>Example for 2 match sets where we 1) want to match in-situ rrs to satellite Rrs and 2) want to match insitu chla to multiple satellite estimates:<br>'rrs_412:Rrs_412;chla_hlpc:chlor_oc2,chlor_oc3,chlor_oc4'<br>The first satellite variable in this list is the 'reference' variable on which the filter criteria are applied. You should always specify an Rrs band rather than a derived product for this first variable otherwise you may limit the returned matchups due to reduced coverage in the derived products. | string | |

| | | | |
|---|---|---|---|
| ***mdb_region*** | MDB region in degrees, format 'lon_min,lon_max,lat_min,lat_max' | any in [-180,180], [-90,90] | -180.0,180.0,-90.0,90.0 |
| ***mdb_start_date*** | MDB start date in format 'yyyymmdd' | yyyymmdd | today |
| ***mdb_start_time*** | MDB start time in format 'hh:mm:ss' | hh:mm:ss | 00:00:00 |
| ***mdb_end_date*** | MDB end date in format 'yyyymmdd' | yyyymmdd | today |
| ***mdb_end_time*** | MDB end time in format 'hh:mm:ss' | hh:mm:ss | 23:59:59 |
| ***time_series_extraction_mode*** | If set to True, time series extraction mode is applied (i.e., start/end times are ignored) | True, False | True |
| ***time_delta*** | Minifiles are extracted from satellite products which cover in situ time +- time delta in hours | Integer > 0 | 1 |
| ***brdf*** | Apply BRDF normalisation (only level 2 products) | True, False | False |
| ***macro_pixel_size*** | Size of macro pixel in minifiles | 3 or 5 | 3 |
| ***sza_max_valid*** | Valid sun zenith angle upper limit in degrees | any in [0,90] | 70.0 |
| ***vza_max_valid*** | Valid sensor zenith angle upper limit in degrees | any in [0,90] | 60.0 |
| ***valid_pixel_expr*** | Valid expression for satellite observation, to be given in Python syntax and dependent on type of Online Processor OC product (e.g., 'not bitmask.LAND' for a Polymer product) | | if left blank, default expression for given L2 processor is used (see Table 3-12) |
| ***min_valid_pixels*** | Minimum number of valid pixels within macro pixel in percent. Default is 50% + 1 pixel. | any in [0,100] | 50.0 |
| ***variance_factor*** | Pixel outliers will be removed if value is outside this | float > 0.0 | 1.5 |
| ***coeff_of_variation_thresh*** | Matchup will be discarded if CV = stdev/mean > threshold | float > 0.0 | 0.2 |

*Table 3-12: Default satellite valid pixel expressions used in OMAPS MATCHUP Module*

| PROCESSOR | DEFAULT VALID EXPRESSION |
|---|---|
| OLCI L1 | `not quality_flags.invalid` |
| OLCI L2 Standard (IPF) | `WQSF.WATER and not WQSF.CLOUD and not WQSF.CLOUD_AMBIGUOUS and not WQSF.CLOUD_MARGIN and not WQSF.INVALID and not WQSF.COSMETIC and not WQSF.SATURATED and not WQSF.SUSPECT and not WQSF.HISOLZEN and not WQSF.HIGHGLINT and not WQSF.SNOW_ICE and not WQSF.AC_FAIL and not WQSF.WHITECAPS and not WQSF.ADJAC and not WQSF.RWNEG_O2 and not WQSF.RWNEG_O3 and not WQSF.RWNEG_O4 and not WQSF.RWNEG_O5 and not WQSF.RWNEG_O6 and not WQSF.RWNEG_O7 and not WQSF.RWNEG_O8` |
| POLYMER | `not bitmask.LAND and not bitmask.CLOUD_BASE and not bitmask.L1_INVALID and not bitmask.NEGATIVE_BB and not bitmask.OUT_OF_BOUNDS and not bitmask.EXCEPTION and not bitmask.THICK_AEROSOL and not` |

| PROCESSOR | DEFAULT VALID EXPRESSION |
|---|---|
| | `bitmask.HIGH_AIR_MASS and not bitmask.EXTERNAL_MASK and not bitmask.CASE2 and not bitmask.INCONSISTENCY` |
| SACSO | `not flags.BAD` |
| L2GEN | `not l2_flags.ATMFAIL and not l2_flags.LAND and not l2_flags.CLDICE and not l2_flags.SEAICE and not l2_flags.HIGLINT and not l2_flags.HILT and not l2_flags.HISATZEN and not l2_flags.STRAYLIGHT and not l2_flags.HISOLZEN and not l2_flags.LOWLW and not l2_flags.CHLFAIL and not l2_flags.ABSAER and not l2_flags.MODGLINT` |

```
[MDB_GENERATION]
; configuration for mdb generation
satellite_input_dir = /data/omaps/omaps_sum_test/testdata/pml_l2_poly
extraction_dir = /data/omaps/omaps_sum_test/testdata/extractions
mdb_output_dir = /data/omaps/omaps_sum_test/testdata/mdb
ocdb_username=olaf
ocdb_password=olaf
processing_label = PML_L2_POLY_TEST
matchup_variables = rrs_412:Rrs_412;rrs_443:Rrs_443;rrs_490:Rrs_490;rrs_510:Rrs_510
mdb_region = -160.0,-140.0,10.0,30.0
mdb_start_date = 20170101
mdb_start_time = 00:00:00
mdb_end_date = 20170331
mdb_end_time = 23:59:59
time_delta = 12
sza_max_valid = 70.0
vza_max_valid = 60.0
valid_pixel_expr =
min_valid_pixels = 40.0
variance_factor = 1.5
coeff_of_variation_thresh = 0.5
macro_pixel_size = 5
sensors = OLCI
platforms = A
processing_level = 2
l2_processor = POLYMER
timeliness = NT
resolution = F
brdf = False
matchup_write_netcdf = True
matchup_write_csv = True
matchup_write_owt = True
owt_normalisation = False
time_series_extraction_mode = False
log_dir = /data/omaps/omaps_sum_test/testdata/log
```

*Figure 6: Example of a config ini file for an OMAPS matchup generation.*

#### *3.2.1.3 Intermediate Data*

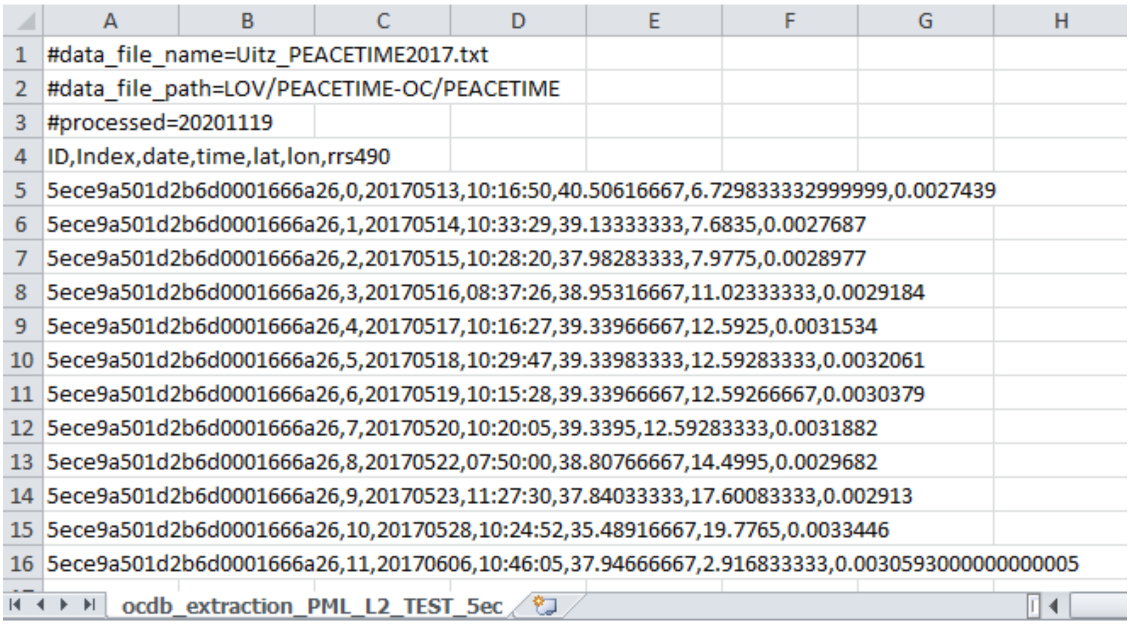
##### 3.2.1.3.1 In situ extraction files

As first step in the matchup generation procedure, in situ data records are extracted from the OCDB, applying the matchup criteria in the configuration file. These records are temporarily stored in a set of CSV files, one per OCDB dataset with contributing in-situ records for the given matchup criteria. An example is shown in Figure 7. The combination of these in situ extraction files with the satellite minifiles (see below) allows a simple creation of the resulting matchups and avoids multiple access to the OCDB and repeated searches in it.

After successful matchup generation, the in-situ extraction files are usually no longer needed and can be deleted.

*Table 3-13: Main properties of an in-situ extraction file*

| ***Processing step*** | **MDB generation** |
|---|---|
| ***Type of product*** | Intermediate product |
| ***Naming convention*** | ocdb_extraction_[LABEL]_[OCDB_DATASET_ID].csv<br>e.g.,<br>'ocdb_extraction_FIRSTTEST_5ece9a501d2b6d0001666a26.csv' |
| ***Storage*** | temporal on disk in subdirectory <extraction_dir>/ocdb/<processing_label> |
| ***Format*** | CSV |
| ***Temporal Coverage*** | variable |
| ***Spatial Coverage*** | variable |
| ***Dimensions*** | variable |
| ***Size per file*** | variable |
| ***Number of files*** | 1 per OCDB dataset with contributing in situ records for given matchup criteria |



| | A |
|---|---|
| 1 | #data_file_name=Uitz_PEACETIME2017.txt |
| 2 | #data_file_path=LOV/PEACETIME-OC/PEACETIME |
| 3 | #processed=20201119 |
| 4 | ID,Index,date,time,lat,lon,rrs490 |
| 5 | 5ece9a501d2b6d0001666a26,0,20170513,10:16:50,40.50616667,6.729833332999999,0.0027439 |
| 6 | 5ece9a501d2b6d0001666a26,1,20170514,10:33:29,39.13333333,7.6835,0.0027687 |
| 7 | 5ece9a501d2b6d0001666a26,2,20170515,10:28:20,37.98283333,7.9775,0.0028977 |
| 8 | 5ece9a501d2b6d0001666a26,3,20170516,08:37:26,38.95316667,11.02333333,0.0029184 |
| 9 | 5ece9a501d2b6d0001666a26,4,20170517,10:16:27,39.33966667,12.5925,0.0031534 |
| 10 | 5ece9a501d2b6d0001666a26,5,20170518,10:29:47,39.33983333,12.59283333,0.0032061 |
| 11 | 5ece9a501d2b6d0001666a26,6,20170519,10:15:28,39.33966667,12.59266667,0.0030379 |
| 12 | 5ece9a501d2b6d0001666a26,7,20170520,10:20:05,39.3395,12.59283333,0.0031882 |
| 13 | 5ece9a501d2b6d0001666a26,8,20170522,07:50:00,38.80766667,14.4995,0.0029682 |
| 14 | 5ece9a501d2b6d0001666a26,9,20170523,11:27:30,37.84033333,17.60083333,0.002913 |
| 15 | 5ece9a501d2b6d0001666a26,10,20170528,10:24:52,35.48916667,19.7765,0.0033446 |
| 16 | 5ece9a501d2b6d0001666a26,11,20170606,10:46:05,37.94666667,2.916833333,0.003059300000000005 |

ocdb_extraction_PML_L2_TEST_5ec

*Figure 7: Example of an in-situ extraction CSV file.*

#### 3.2.1.3.2 Satellite minifiles

For each in situ data record, a corresponding satellite 'minifile', holding a $N_{macro}$x$N_{macro}$ macro pixel window around the selected site (in situ lat, lon) for the matching time range, is extracted from the available satellite input data archive. This means, of course, that one satellite extraction can correspond to several in situ data records.

The main properties of a satellite extraction minifile are listed in Table 3-14. The product format and content are the same as for the full satellite input products. An example is shown in Figure 8, where (preliminary) Ocean Colour Processor products were used as input.

After successful matchup generation, the satellite minifiles can be deleted as well. However, they might still be useful for other purposes.

*Table 3-14: Main properties of a satellite extraction minifile*

| ***Processing step*** | **MDB generation** |
|---|---|
| ***Type of product*** | Intermediate product |
| ***Naming convention*** | [SATELLITE_PRODUCT_NAME].[E\|W]DDDddd[N\|S]DDDddd_<$N_{macro}$>x<$N_{macro}$>.nc<br>e.g.,<br>'S3A_OL_2_WFR____20170522T090453_20170522T090653_20171018T123841_0119_018_050______MR1_R_NT_002.E035205N050102_5x5.nc'<br><br>where "E035205N050102" means that the centroid of the macropixel is located in (50.102N,35.205E). |
| ***Storage*** | temporal on disk in subdirectory<br><extraction_dir>/sat_minifiles/<processing_label> |
| ***Format*** | NetCDF |
| ***Temporal Coverage*** | same as full satellite product |
| ***Spatial Coverage*** | $N_{macro}$ x$N_{macro}$ pixel |
| ***Dimensions*** | $N_{macro}$ x$N_{macro}$ |
| ***Size per file*** | ~ 50KB |
| ***Number of files*** | max. 1 per in situ data record for given matchup criteria (one satellite extraction can correspond to several in situ data records) |

```
netcdf S3A_OL_2_WFR____20170522T090453_20170522T090653_20171018T123841_0119_018_050______MR1_R_NT_002.nc.extract_5x5 {
dimensions:
        rows = 5 ;
        columns = 5 ;
variables:
        double Rrs_412(rows, columns) ;
                Rrs_412:description = "water reflectance (dimensionless ; fully normalized)" ;
        double Rrs_443(rows, columns) ;
                Rrs_443:description = "water reflectance (dimensionless ; fully normalized)" ;
        double Rrs_490(rows, columns) ;
                Rrs_490:description = "water reflectance (dimensionless ; fully normalized)" ;
        double Rrs_510(rows, columns) ;
                Rrs_510:description = "water reflectance (dimensionless ; fully normalized)" ;
        double Rrs_560(rows, columns) ;
                Rrs_560:description = "water reflectance (dimensionless ; fully normalized)" ;
        double Rrs_665(rows, columns) ;
                Rrs_665:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw400(rows, columns) ;
                Rw400:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw412(rows, columns) ;
                Rw412:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw443(rows, columns) ;
                Rw443:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw490(rows, columns) ;
                Rw490:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw510(rows, columns) ;
                Rw510:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw560(rows, columns) ;
                Rw560:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw620(rows, columns) ;
                Rw620:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw665(rows, columns) ;
                Rw665:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw681(rows, columns) ;
                Rw681:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw709(rows, columns) ;
                Rw709:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw754(rows, columns) ;
                Rw754:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw779(rows, columns) ;
                Rw779:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw865(rows, columns) ;
                Rw865:description = "water reflectance (dimensionless ; fully normalized)" ;
        float Rw1020(rows, columns) ;
                Rw1020:description = "water reflectance (dimensionless ; fully normalized)" ;
        float logchl(rows, columns) ;
                logchl:description = "log10 of the chl-a concentration in mg/m3" ;
        float Rnir(rows, columns) ;
        float Rgli(rows, columns) ;
        float bbs(rows, columns) ;
        short bitmask(rows, columns) ;
                bitmask:description = "LAND:1, CLOUD_BASE:2, L1_INVALID:4, NEGATIVE_BB:8, OUT_OF_BOUNDS:16, EXCEPTION:32
, THICK_AEROSOL:64, HIGH_AIR_MASS:128, EXTERNAL_MASK:512, CASE2:1024, INCONSISTENCY:2048" ;
                bitmask:bitmask_reject = "bitmask & 1023 != 0" ;
        double latitude(rows, columns) ;
        double longitude(rows, columns) ;

// global attributes:
                :title = "EUMETSAT Sentinel 3, OLCI Full Resolution, Extraction Subset Data." ;
                :institution = "EUropean METeorological SATellite organisation (EUMETSAT)." ;
                :contact = "ops@eumetsat.int" ;
                :creation = "2020-10-01T23:11:07.661261" ;
                :centre = "MAR" ;
                :comment = "Original OLCI full resolution products imported and subset to a small grid, with tie_point d
ata interpolated to full resolution, and compiled into a single file. " ;
                :version = "0.5" ;
```

*Figure 8: Example of a satellite 5x5 pixel minifile (ncdump output).*

### 3.2.1.4 Output Data

Matchups Database (MDB) products are netCDF files which include the matchups between the satellite data (OLCI) and the in-situ data from Copernicus OCDB for the specified matchup criteria. As said earlier, Level-1 or Level-2 matchups can be generated in OMAPS, to be specified in the config file. As an option, the user can additionally (or exclusively) create a MDB ‘summary file’ in CSV format, which contains most of the content of the netCDF product in tabular form, as described further below.

Each MDB netCDF file contains the following information:

- for all matching in situ records ($N_{insitu}$) and all corresponding satellite macropixels ($N_{satellite}$), the matchup variables specified in the configuration, obtained from in situ
- for all pixels (rows, columns) of all satellite macropixels ($N_{satellite}$), the matchup variables specified in the configuration, obtained from satellite

- for all satellite macropixels ($N_{satellite}$), the macropixel mean value of each matchup variable specified in the configuration
- for all satellite macropixels ($N_{satellite}$), the number of valid pixels in the micropixel for the 'reference' matchup variable (the variable on which the filter criteria are applied) specified in the configuration
- optionally and if available, for all pixels (rows, columns) of all satellite macropixels ($N_{satellite}$), water classes obtained from satellite (depending on OC product type and algorithm)
- if available, for all pixels (rows, columns) of all satellite macropixels ($N_{satellite}$), the satellite flag variables (depending on OC product type and algorithm)
- for all satellite macropixels ($N_{satellite}$), the central latitude, longitude, and time
- for all matching in situ records ($N_{insitu}$), the OCDB extraction ID
- for all matching in situ records ($N_{insitu}$) and all corresponding satellite macropixels ($N_{satellite}$), the in-situ latitude, longitude, and time
- for all matching in situ records ($N_{insitu}$) and all corresponding satellite macropixels ($N_{satellite}$), the time difference between satellite and in situ measurement in seconds
- for all in situ / satellite matchup variables in Level-2 matchups, the Median Absolute Difference (MdAD) of all matchups in the MDB product
- for all in situ / satellite matchup variables in Level-2 matchups, the Median Absolute Percentage Difference (MdAPD) of all matchups in the MDB product
- if Rrs variables are present in the Level-2 matchups, the Spectral Angle Mapper (SAM) for all in situ / satellite Rrs variables at given wavelengths, and for all matchups in the MDB product
- if Rrs variables are present in the Level-2 matchups, the Chi-Square norm for all in situ / satellite Rrs variables at given wavelengths, and for all matchups in the MDB product

All variables are described in more detail in Table 3-17. The statistical quantities (MdAD, MdAPD, SAM, CHI_SQUARE) are not stored as variables, but as global attributes.

Each MDB 'summary' CSV file contains almost the same information, except that the satellite micropixel mean values are not written into this file. An additional column 'pixel_ID' is added to introduce a unique ID for the assignment of each single satellite measurement and each in situ record within an OCDB dataset. Moreover, for all satellite matchup variables, both filtered and original data are written in separate columns. The reason for keeping the original, unfiltered data here is their potential usage in the RR-AC Module. The statistical quantities (MdAD, MdAPD, SAM, CHI_SQUARE) are written into the CSV header.

The main properties of these MDB products are listed in Table 3-15. Each MDB file is named using the following template: MDB_S3[PLATFORM]_[SENSOR]_[LEVEL]_[PROCESSING_LABEL].[EXTENSION]. The meaning of the fields composing this file name is explained in Table 3-16. The basic content of a MDB netCDF product is listed in Table 3-17, an example is shown in Figure 9 to Figure 11, where Online Processor products originating from Polymer were used as input for the matchup generation. Snippets from the corresponding CSV summary files are shown in Figure 12 and Figure 13.

*Table 3-15: Main properties of a MDB product*

| | |
|---|---|
| ***Processing step*** | **MDB generation** |
| ***Type of product*** | MDB final product |
| ***Naming convention*** | MDB_S3[PLATFORM]_[SENSOR]_[LEVEL]_[PROCESSING_LABEL].[EXTENSION]<br>e.g., *MDB_S3A_OLCI_L2_FIRSTTEST.nc* or *MDB_S3A_OLCI_L2_FIRSTTEST.csv* |
| ***Storage*** | tbd |
| ***Format*** | NetCDF (standard MDB) and/or CSV ('summary' file) |
| ***Temporal Coverage*** | variable |
| ***Spatial Coverage*** | variable |
| ***Dimensions*** | satellite variables:<br>$N_{macro}$ x $N_{macro}$ x $N_{satellite}$ (macropixel size x number of matching minifiles)<br><br>in situ variables:<br>$N_{insitu}$ x $N_{satellite}$ (number of matching in situ records x number of matching minifiles) |
| ***Size per file*** | variable |
| ***Number of files*** | variable (1 per matchup processing with a given configuration) |

*Table 3-16: MDB product: File name convention details*

| ***Field ID*** | ***Meaning*** | ***Possible values*** | ***Default*** |
|---|---|---|---|
| **PLATFORM** | OLCI A or B | A, B | none |
| **SENSOR** | Sensor | OLCI | OLCI |
| **LEVEL** | OLCI products level: | 'L1' for EFR products or 'L2' for WFR products | none |
| **PROCESSING_LABEL** | Label for identification of processing runs | arbitrary string<br>(label is used in filename, thus use allowed characters only) | none |
| **EXTENSION** | in-situ data type | '.nc' or '.csv' | none |
| **"_"** | Separation character | | |

*Table 3-17: Dimensions in a MDB netCDF product*

| DIMENSION | VALUE RANGE | DESCRIPTION |
|---|---|---|
| rows | 3 or 5 | x size of macropixel |
| columns | 3 or 5 | y size of macropixel |
| satellite_id | UNLIMITED | number of satellite extractions ('minifiles') in given MDB product |

| DIMENSION | VALUE RANGE | DESCRIPTION |
| --- | --- | --- |
| insitu_id | UNLIMITED | number of insitu measurements in given MDB product |
| wind_vectors | 2 | Taken from OLCI L1 (in Level-1 matchups only) |
| satellite_detectors | 3700 | Taken from OLCI L1 (in Level-1 matchups only) |
| satellite_bands | 21 | Taken from OLCI L1 (in Level-1 matchups only) |

*Table 3-18: Variables in a MDB netCDF product*

| PARAMETER | UNIT | DATA TYPE | SIZE (BYTES) | DESCRIPTION |
| --- | --- | --- | --- | --- |
| insitu_<variable> | variable | double | 8 | matching in situ variable as specified in the configuration |
| satellite_<variable> | variable | double | 8 | matching satellite variable as specified in the configuration |
| satellite_num_valid_pixels _<reference variable> | none | int | 2 | number of valid pixels in the micropixel for the 'reference' matchup variable (the variable on which the filter criteria are applied) specified in the configuration |
| satellite_<flags> | none | byte or int | 1/2 | satellite flag variables (depending on OC product type and algorithm). E.g., 'satellite_bitmask' in case of Polymer products. |
| water_class<n> (optional) | none | double | 8 | water classes obtained from satellite (depending on OC product type and algorithm). E.g., n=1,..,14 for Polymer products. Optional output (user defined). |
| satellite_latitude | deg | double | 8 | satellite per-pixel latitude |
| satellite_longitude | deg | double | 8 | satellite per-pixel longitude |
| central_latitude | degrees_north | float | 8 | macropixel central latitude |
| central_longitude | degrees_east | float | 8 | macropixel central longitude |
| central_time | date | char | 19 | macropixel central time in yyyy-MM-ddThh:mm:ss |
| insitu_lat | degrees_north | float | 8 | latitude of in situ measurement |
| insitu_lon | degrees_east | float | 8 | longitude of in situ measurement |
| insitu_time | date | char | 19 | time of in situ measurement in yyyy-MM-ddThh:mm:ss |

| PARAMETER | UNIT | DATA TYPE | SIZE (BYTES) | DESCRIPTION |
|---|---|---|---|---|
| insitu_ocdb_extraction_id | none | String | variable | OCDB extraction IDs for the in-situ records |
| time_difference | sec | int | 4 | time difference of in situ measurement and macropixel central time in seconds |

*Table 3-19: Global attributes in a MDB netCDF product*

| ATTRIBUTE | DESCRIPTION |
|---|---|
| creation_time | matching in situ variable as specified in the configuration |
| satellite | matching satellite variable as specified in the configuration |
| sensor | number of valid pixels in the micropixel for the 'reference' matchup variable (the variable on which the filter criteria are applied) specified in the configuration |
| description | satellite flag variables (depending on OC product type and algorithm). E.g., 'satellite_bitmask' in case of Polymer products. |
| AC_processor | water classes obtained from satellite (depending on OC product type and algorithm). E.g., n=1,..,14 for Polymer products. |
| AC_processor_version | satellite per-pixel latitude |
| macro_pixel_size | satellite per-pixel longitude |
| sat_bands | macropixel central latitude |
| sat_variables | macropixel central longitude |
| insitu_variables | macropixel central time in yyyy-MM-ddThh:mm:ss |
| MdAD | for all matchup variables, Median Absolute Difference (MdAD) of all matchups in the MDB product. Given as semicolon-separated string and for Level-2 matchups only. |
| MdAPD | for given matchup variables, Median Absolute Percentage Difference (MdAD) of all matchups in the MDB product. Given as semicolon-separated string and for Level-2 matchups only. |
| SAM | for given Rrs matchup variables, Spectral Angle Mapper (SAM) of all matchups in the MDB product. Given for Level-2 matchups only. |
| CHI_SQR | for given Rrs matchup variables, the Chi-Square norm of all matchups in the MDB product. Given for Level-2 matchups only. |

```
netcdf MDB_S3A_OLCI_L2_OCDB_PML_L2_RRS {
dimensions:
        rows = 5 ;
        columns = 5 ;
        satellite_id = UNLIMITED ; // (3 currently)
        insitu_id = UNLIMITED ; // (4 currently)
variables:
        int satellite_num_valid_pixels_Rrs_412(satellite_id) ;
        double satellite_Rrs_412(satellite_id, rows, columns) ;
                satellite_Rrs_412:long_name = "water reflectance (dimensionless ; fully normalised)" ;
                satellite_Rrs_412:units = "sr-1" ;
        double satellite_Rrs_412_mean(satellite_id) ;
                satellite_Rrs_412_mean:long_name = "water reflectance, mean of macropixel (dimensionless ; fully normalised)" ;
                satellite_Rrs_412_mean:units = "-" ;
        double satellite_Rrs_443(satellite_id, rows, columns) ;
                satellite_Rrs_443:long_name = "water reflectance (dimensionless ; fully normalised)" ;
                satellite_Rrs_443:units = "sr-1" ;
        double satellite_Rrs_443_mean(satellite_id) ;
                satellite_Rrs_443_mean:long_name = "water reflectance, mean of macropixel (dimensionless ; fully normalised)" ;
                satellite_Rrs_443_mean:units = "-" ;
        double satellite_Rrs_490(satellite_id, rows, columns) ;
                satellite_Rrs_490:long_name = "water reflectance (dimensionless ; fully normalised)" ;
                satellite_Rrs_490:units = "sr-1" ;
        double satellite_Rrs_490_mean(satellite_id) ;
                satellite_Rrs_490_mean:long_name = "water reflectance, mean of macropixel (dimensionless ; fully normalised)" ;
                satellite_Rrs_490_mean:units = "-" ;
        double satellite_Rrs_510(satellite_id, rows, columns) ;
                satellite_Rrs_510:long_name = "water reflectance (dimensionless ; fully normalised)" ;
                satellite_Rrs_510:units = "sr-1" ;
        double satellite_Rrs_510_mean(satellite_id) ;
                satellite_Rrs_510_mean:long_name = "water reflectance, mean of macropixel (dimensionless ; fully normalised)" ;
                satellite_Rrs_510_mean:units = "-" ;
        double satellite_Rrs_560(satellite_id, rows, columns) ;
                satellite_Rrs_560:long_name = "water reflectance (dimensionless ; fully normalised)" ;
                satellite_Rrs_560:units = "sr-1" ;
        double satellite_Rrs_560_mean(satellite_id) ;
                satellite_Rrs_560_mean:long_name = "water reflectance, mean of macropixel (dimensionless ; fully normalised)" ;
                satellite_Rrs_560_mean:units = "-" ;
        double satellite_Rrs_620(satellite_id, rows, columns) ;
                satellite_Rrs_620:long_name = "water reflectance (dimensionless ; fully normalised)" ;
                satellite_Rrs_620:units = "sr-1" ;
        double satellite_Rrs_620_mean(satellite_id) ;
                satellite_Rrs_620_mean:long_name = "water reflectance, mean of macropixel (dimensionless ; fully normalised)" ;
                satellite_Rrs_620_mean:units = "-" ;
        double satellite_Rrs_665(satellite_id, rows, columns) ;
                satellite_Rrs_665:long_name = "water reflectance (dimensionless ; fully normalised)" ;
                satellite_Rrs_665:units = "sr-1" ;
        double satellite_Rrs_665_mean(satellite_id) ;
                satellite_Rrs_665_mean:long_name = "water reflectance, mean of macropixel (dimensionless ; fully normalised)" ;
                satellite_Rrs_665_mean:units = "-" ;
        double satellite_Rrs_865(satellite_id, rows, columns) ;
                satellite_Rrs_865:long_name = "water reflectance (dimensionless ; fully normalised)" ;
                satellite_Rrs_865:units = "sr-1" ;
        double satellite_Rrs_865_mean(satellite_id) ;
                satellite_Rrs_865_mean:long_name = "water reflectance, mean of macropixel (dimensionless ; fully normalised)" ;
                satellite_Rrs_865_mean:units = "-" ;
        double satellite_Rrs_1020(satellite_id, rows, columns) ;
                satellite_Rrs_1020:long_name = "water reflectance (dimensionless ; fully normalised)" ;
                satellite_Rrs_1020:units = "sr-1" ;
        double satellite_Rrs_1020_mean(satellite_id) ;
                satellite_Rrs_1020_mean:long_name = "water reflectance, mean of macropixel (dimensionless ; fully normalised)" ;
                satellite_Rrs_1020_mean:units = "-" ;
        byte satellite_bitmask(satellite_id, rows, columns) ;
                satellite_bitmask:description = "LAND:1, CLOUD_BASE:2, L1_INVALID:4, NEGATIVE_BB:8, OUT_OF_BOUNDS:16, EXCEPTION:32,
4, HIGH_AIR_MASS:128, EXTERNAL_MASK:512, CASE2:1024, INCONSISTENCY:2048" ;
                satellite_bitmask:bitmask_reject = "bitmask & 1023 != 0" ;
        double satellite_latitude(satellite_id, rows, columns) ;
        double satellite_longitude(satellite_id, rows, columns) ;
```

*Figure 9: Example of a MDB NetCDF product (ncdump output, part 1).*

```
	double water_class1(satellite_id, rows, columns) ;
		water_class1:_FillValue = NaN ;
	double water_class2(satellite_id, rows, columns) ;
		water_class2:_FillValue = NaN ;
	double water_class3(satellite_id, rows, columns) ;
		water_class3:_FillValue = NaN ;
	double water_class4(satellite_id, rows, columns) ;
		water_class4:_FillValue = NaN ;
	double water_class5(satellite_id, rows, columns) ;
		water_class5:_FillValue = NaN ;
	double water_class6(satellite_id, rows, columns) ;
		water_class6:_FillValue = NaN ;
	double water_class7(satellite_id, rows, columns) ;
		water_class7:_FillValue = NaN ;
	double water_class8(satellite_id, rows, columns) ;
		water_class8:_FillValue = NaN ;
	double water_class9(satellite_id, rows, columns) ;
		water_class9:_FillValue = NaN ;
	double water_class10(satellite_id, rows, columns) ;
		water_class10:_FillValue = NaN ;
	double water_class11(satellite_id, rows, columns) ;
		water_class11:_FillValue = NaN ;
	double water_class12(satellite_id, rows, columns) ;
		water_class12:_FillValue = NaN ;
	double water_class13(satellite_id, rows, columns) ;
		water_class13:_FillValue = NaN ;
	double water_class14(satellite_id, rows, columns) ;
		water_class14:_FillValue = NaN ;
	double central_latitude(satellite_id) ;
		central_latitude:long_name = "Central pixel latitude " ;
		central_latitude:units = "degrees_north " ;
	double central_longitude(satellite_id) ;
		central_longitude:long_name = "Central pixel longitude " ;
		central_longitude:units = "degrees_east " ;
	string central_time(satellite_id) ;
		central_time:long_name = "Central pixel time " ;
	float insitu_rrs_412(satellite_id, insitu_id) ;
		insitu_rrs_412:description = "" ;
		insitu_rrs_412:units = "" ;
	float insitu_rrs_443(satellite_id, insitu_id) ;
		insitu_rrs_443:description = "" ;
		insitu_rrs_443:units = "" ;
	float insitu_rrs_490(satellite_id, insitu_id) ;
		insitu_rrs_490:description = "" ;
		insitu_rrs_490:units = "" ;
	float insitu_rrs_510(satellite_id, insitu_id) ;
		insitu_rrs_510:description = "" ;
		insitu_rrs_510:units = "" ;
	float insitu_rrs_560.5796(satellite_id, insitu_id) ;
		insitu_rrs_560.5796:description = "" ;
		insitu_rrs_560.5796:units = "" ;
	float insitu_rrs_620.626(satellite_id, insitu_id) ;
		insitu_rrs_620.626:description = "" ;
		insitu_rrs_620.626:units = "" ;
	float insitu_rrs_665(satellite_id, insitu_id) ;
		insitu_rrs_665:description = "" ;
		insitu_rrs_665:units = "" ;
	float insitu_rrs_870(satellite_id, insitu_id) ;
		insitu_rrs_870:description = "" ;
		insitu_rrs_870:units = "" ;
	float insitu_rrs_1020(satellite_id, insitu_id) ;
		insitu_rrs_1020:description = "" ;
		insitu_rrs_1020:units = "" ;
```

*Figure 10: Example of a MDB NetCDF product (ncdump output, part 2).*

```
	string insitu_time(satellite_id, insitu_id) ;
		insitu_time:long_name = "Insitu time " ;
	float insitu_lat(satellite_id, insitu_id) ;
		insitu_lat:long_name = "Insitu latitude " ;
		insitu_lat:units = "degrees_north " ;
	float insitu_lon(satellite_id, insitu_id) ;
		insitu_lon:long_name = "Insitu longitude " ;
		insitu_lon:units = "degrees_east " ;
	int time_difference(satellite_id, insitu_id) ;
		time_difference:long_name = "Satellite/Insitu time difference " ;
		time_difference:units = "seconds " ;
	string insitu_ocdb_extraction_id(satellite_id, insitu_id) ;
		insitu_ocdb_extraction_id:long_name = "Insitu data set ID in OCDB " ;

// global attributes:
		:creation_time = "2021-07-28T14:53:18.252964" ;
		:satellite = "S3" ;
		:sensor = "OLCI" ;
		:description = "S3 OLCI - OCDB Matchups Data Base" ;
		:data_policy = "tbd" ;
		:AC_processor = "polymer" ;
		:AC_processor_version = "4.17" ;
		:macro_pixel_size = "5" ;
		:sat_bands = "Rrs_400;Rrs_412;Rrs_443;Rrs_490;Rrs_510;Rrs_560;Rrs_620;Rrs_665;Rrs_681;Rrs_709;Rrs_754;Rrs_779;Rrs_865;Rrs_1020"
		:sat_variables = "Rrs_412;Rrs_443;Rrs_490;Rrs_510;Rrs_560;Rrs_620;Rrs_665;Rrs_865;Rrs_1020" ;
		:insitu_variables = "rrs_412;rrs_443;rrs_490;rrs_510;rrs_560.5796;rrs_620.626;rrs_665;rrs_870;rrs_1020" ;
		:MdAD = "0.001095;0.000770;0.000351;0.000377;0.000058;0.000022;0.000033;nan;nan" ;
		:MdAPD = "8.712149;7.851668;5.863235;10.895792;4.785553;12.076043;34.209702;nan;nan" ;
		:SAM = "0.018194" ;
		:CHI_SQUARE = "0.085788" ;
}
```

*Figure 11: Example of a MDB NetCDF product (ncdump output, part 3).*

| | A | B | C | D | E | F | G | H | I | J | K | L | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | #creation_time=2021-07-28T00:18:03.422494 | | | | | | | | | | | | |
| 2 | #satellite=S3 | | | | | | | | | | | | |
| 3 | #sensor=OLCI | | | | | | | | | | | | |
| 4 | #description=S3 OLCI - OCDB Matchups Data Base | | | | | | | | | | | | |
| 5 | #data_policy=tbd | | | | | | | | | | | | |
| 6 | #AC_processor=polymer | | | | | | | | | | | | |
| 7 | #AC_processor_version=4.17 | | | | | | | | | | | | |
| 8 | #macro_pixel_size=5 | | | | | | | | | | | | |
| 9 | #sat_bands=Rrs_400;Rrs_412;Rrs_443;Rrs_490;Rrs_510;Rrs_560;Rrs_620;Rrs_665;Rrs_681;Rrs_709;Rrs_754;Rrs_779;Rrs_865;Rrs_1020 | | | | | | | | | | | | |
| 10 | #sat_variables=Rrs_412;Rrs_443;Rrs_490;Rrs_510;Rrs_560;Rrs_620;Rrs_665;Rrs_865;Rrs_1020 | | | | | | | | | | | | |
| 11 | #insitu_variables=rrs_412;rrs_443;rrs_490;rrs_510;rrs_560.5796;rrs_620.626;rrs_665;rrs_870;rrs_1020 | | | | | | | | | | | | |
| 12 | #MDAD=0.001095;0.000770;0.000351;0.000377;0.000058;0.000022;0.000033;nan;nan | | | | | | | | | | | | |
| 13 | #MDAPD=8.712149;7.851668;5.863235;10.895792;4.785553;12.076043;34.209702;nan;nan | | | | | | | | | | | | |
| 14 | #SAM=0.018194 | | | | | | | | | | | | |
| 15 | #CHI_SQUARE=0.085788 | | | | | | | | | | | | |
| 16 | pixel_ID | central_latitude | central_longitude | central_time | satellite_latitude | satellite_longitude | insitu_lat | insitu_lon | insitu_time | time_difference | insitu_rrs_412 | insitu_rrs_443 | insitu_ |
| 17 | S3A_OL_1_EFR____20170223T203804_0_0_0-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8145 | -157.1939 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 18 | S3A_OL_1_EFR____20170223T203804_0_0_1-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8140 | -157.1913 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 19 | S3A_OL_1_EFR____20170223T203804_0_0_2-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8135 | -157.1888 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 20 | S3A_OL_1_EFR____20170223T203804_0_0_3-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8129 | -157.1863 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 21 | S3A_OL_1_EFR____20170223T203804_0_0_4-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8124 | -157.1838 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 22 | S3A_OL_1_EFR____20170223T203804_0_1_0-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8119 | -157.1945 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 23 | S3A_OL_1_EFR____20170223T203804_0_1_1-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8114 | -157.1920 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 24 | S3A_OL_1_EFR____20170223T203804_0_1_2-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8109 | -157.1894 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 25 | S3A_OL_1_EFR____20170223T203804_0_1_3-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8103 | -157.1869 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 26 | S3A_OL_1_EFR____20170223T203804_0_1_4-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8098 | -157.1844 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 27 | S3A_OL_1_EFR____20170223T203804_0_2_0-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8093 | -157.1951 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 28 | S3A_OL_1_EFR____20170223T203804_0_2_1-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8088 | -157.1926 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 29 | S3A_OL_1_EFR____20170223T203804_0_2_2-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8083 | -157.1900 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 30 | S3A_OL_1_EFR____20170223T203804_0_2_3-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8077 | -157.1875 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 31 | S3A_OL_1_EFR____20170223T203804_0_2_4-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8072 | -157.1850 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 32 | S3A_OL_1_EFR____20170223T203804_0_3_0-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8067 | -157.1957 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 33 | S3A_OL_1_EFR____20170223T203804_0_3_1-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8062 | -157.1932 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 34 | S3A_OL_1_EFR____20170223T203804_0_3_2-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8057 | -157.1907 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 35 | S3A_OL_1_EFR____20170223T203804_0_3_3-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8051 | -157.1881 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 36 | S3A_OL_1_EFR____20170223T203804_0_3_4-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8046 | -157.1856 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 37 | S3A_OL_1_EFR____20170223T203804_0_4_0-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8041 | -157.1963 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 38 | S3A_OL_1_EFR____20170223T203804_0_4_1-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8036 | -157.1938 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 39 | S3A_OL_1_EFR____20170223T203804_0_4_2-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8031 | -157.1913 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 40 | S3A_OL_1_EFR____20170223T203804_0_4_3-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8025 | -157.1888 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 41 | S3A_OL_1_EFR____20170223T203804_0_4_4-60f6e4bd1d2b6d0001652599_2319 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8020 | -157.1862 | 20.8095 | -157.1905 | 2017-02-23T22:45:00 | 7586 | 0.012312 | 0.009400 | 0.005 |
| 42 | S3A_OL_1_EFR____20170223T203804_1_0_0-60f6e4bd1d2b6d0001652599_2320 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8145 | -157.1939 | 20.8095 | -157.1905 | 2017-02-24T00:22:00 | 13406 | 0.012740 | 0.009692 | 0.006 |
| 43 | S3A_OL_1_EFR____20170223T203804_1_0_1-60f6e4bd1d2b6d0001652599_2320 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8140 | -157.1913 | 20.8095 | -157.1905 | 2017-02-24T00:22:00 | 13406 | 0.012740 | 0.009692 | 0.006 |
| 44 | S3A_OL_1_EFR____20170223T203804_1_0_2-60f6e4bd1d2b6d0001652599_2320 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8135 | -157.1888 | 20.8095 | -157.1905 | 2017-02-24T00:22:00 | 13406 | 0.012740 | 0.009692 | 0.006 |
| 45 | S3A_OL_1_EFR____20170223T203804_1_0_3-60f6e4bd1d2b6d0001652599_2320 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8129 | -157.1863 | 20.8095 | -157.1905 | 2017-02-24T00:22:00 | 13406 | 0.012740 | 0.009692 | 0.006 |
| 46 | S3A_OL_1_EFR____20170223T203804_1_0_4-60f6e4bd1d2b6d0001652599_2320 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8124 | -157.1838 | 20.8095 | -157.1905 | 2017-02-24T00:22:00 | 13406 | 0.012740 | 0.009692 | 0.006 |
| 47 | S3A_OL_1_EFR____20170223T203804_1_1_0-60f6e4bd1d2b6d0001652599_2320 | 20.8083 | -157.1900 | 2017-02-23T20:38:34 | 20.8119 | -157.1945 | 20.8095 | -157.1905 | 2017-02-24T00:22:00 | 13406 | 0.012740 | 0.009692 | 0.006 |

*Figure 12: Example of a MDB CSV summary file (snippet showing header and first few rows/columns).*

| | K | L | M | N | U | V | W | X | Y | Z | AA | AB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 16 | insitu_rrs_412 | insitu_rrs_443 | insitu_rrs_490 | insitu_rrs_510 | satellite_Rrs_412 | satellite_Rrs_412_filtered | satellite_Rrs_443 | satellite_Rrs_443_filtered | satellite_Rrs_490 | satellite_Rrs_490_filtered | satellite_Rrs_510 | satellite_Rrs_510_filtered |
| 17 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011191 | 0.011191 | 0.008696 | 0.008696 | 0.005721 | 0.005721 | 0.003086 | 0.003086 |
| 18 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011168 | 0.011168 | 0.008592 | 0.008592 | 0.005694 | 0.005694 | 0.003054 | 0.003054 |
| 19 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011934 | 0.011934 | 0.009020 | 0.009020 | 0.005856 | 0.005856 | 0.003117 | 0.003117 |
| 20 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.010971 | 0.010971 | 0.008746 | 0.008746 | 0.005647 | 0.005647 | 0.003181 | 0.003181 |
| 21 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011170 | 0.011170 | 0.008697 | 0.008697 | 0.005651 | 0.005651 | 0.003024 | 0.003024 |
| 22 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.010706 | 0.010706 | 0.008546 | 0.008546 | 0.005478 | 0.005478 | 0.003069 | 0.003069 |
| 23 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011549 | 0.011549 | 0.008928 | 0.008928 | 0.005669 | 0.005669 | 0.003000 | 0.003000 |
| 24 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011470 | 0.011470 | 0.008845 | 0.008845 | 0.005489 | 0.005489 | 0.002940 | nan |
| 25 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.010355 | nan | 0.008291 | nan | 0.005266 | nan | 0.002980 | nan |
| 26 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.010894 | 0.010894 | 0.008734 | 0.008734 | 0.005664 | 0.005664 | 0.003089 | 0.003089 |
| 27 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.012425 | nan | 0.009393 | nan | 0.006173 | nan | 0.003391 | nan |
| 28 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011765 | 0.011765 | 0.009246 | 0.009246 | 0.006035 | 0.006035 | 0.003230 | 0.003230 |
| 29 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.012396 | nan | 0.009483 | nan | 0.006101 | nan | 0.003271 | nan |
| 30 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011163 | 0.011163 | 0.008544 | 0.008544 | 0.005535 | 0.005535 | 0.003028 | 0.003028 |
| 31 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011495 | 0.011495 | 0.008984 | 0.008984 | 0.005717 | 0.005717 | 0.003080 | 0.003080 |
| 32 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.012111 | 0.012111 | 0.009135 | 0.009135 | 0.005881 | 0.005881 | 0.003105 | 0.003105 |
| 33 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011518 | 0.011518 | 0.008932 | 0.008932 | 0.005739 | 0.005739 | 0.003089 | 0.003089 |
| 34 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.012223 | 0.012223 | 0.009160 | 0.009160 | 0.005987 | 0.005987 | 0.003270 | nan |
| 35 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011262 | 0.011262 | 0.008963 | 0.008963 | 0.005760 | 0.005760 | 0.003048 | 0.003048 |
| 36 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011434 | 0.011434 | 0.008837 | 0.008837 | 0.005587 | 0.005587 | 0.003033 | 0.003033 |
| 37 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.012331 | nan | 0.009331 | nan | 0.006075 | nan | 0.003216 | nan |
| 38 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.012237 | 0.012237 | 0.009392 | nan | 0.005980 | 0.005980 | 0.003199 | 0.003199 |
| 39 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.010917 | 0.010917 | 0.008558 | 0.008558 | 0.005528 | 0.005528 | 0.002986 | 0.002986 |
| 40 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011175 | 0.011175 | 0.008579 | 0.008579 | 0.005434 | 0.005434 | 0.002992 | 0.002992 |
| 41 | 0.012312 | 0.009400 | 0.005916 | 0.003418 | 0.011703 | 0.011703 | 0.008930 | 0.008930 | 0.005710 | 0.005710 | 0.003095 | 0.003095 |
| 42 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.011191 | 0.011191 | 0.008696 | 0.008696 | 0.005721 | 0.005721 | 0.003086 | 0.003086 |
| 43 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.011168 | 0.011168 | 0.008592 | 0.008592 | 0.005694 | 0.005694 | 0.003054 | 0.003054 |
| 44 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.011934 | 0.011934 | 0.009020 | 0.009020 | 0.005856 | 0.005856 | 0.003117 | 0.003117 |
| 45 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.010971 | 0.010971 | 0.008746 | 0.008746 | 0.005647 | 0.005647 | 0.003181 | 0.003181 |
| 46 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.011170 | 0.011170 | 0.008697 | 0.008697 | 0.005651 | 0.005651 | 0.003024 | 0.003024 |
| 47 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.010706 | 0.010706 | 0.008546 | 0.008546 | 0.005478 | 0.005478 | 0.003069 | 0.003069 |
| 48 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.011549 | 0.011549 | 0.008928 | 0.008928 | 0.005669 | 0.005669 | 0.003000 | 0.003000 |
| 49 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.011470 | 0.011470 | 0.008845 | 0.008845 | 0.005489 | 0.005489 | 0.002940 | nan |
| 50 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.010355 | nan | 0.008291 | nan | 0.005266 | nan | 0.002980 | nan |
| 51 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.010894 | 0.010894 | 0.008734 | 0.008734 | 0.005664 | 0.005664 | 0.003089 | 0.003089 |
| 52 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.012425 | nan | 0.009393 | nan | 0.006173 | nan | 0.003391 | nan |
| 53 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.011765 | 0.011765 | 0.009246 | 0.009246 | 0.006035 | 0.006035 | 0.003230 | 0.003230 |
| 54 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.012396 | nan | 0.009483 | nan | 0.006101 | nan | 0.003271 | nan |
| 55 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.011163 | 0.011163 | 0.008544 | 0.008544 | 0.005535 | 0.005535 | 0.003028 | 0.003028 |
| 56 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.011495 | 0.011495 | 0.008984 | 0.008984 | 0.005717 | 0.005717 | 0.003080 | 0.003080 |
| 57 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.012111 | 0.012111 | 0.009135 | 0.009135 | 0.005881 | 0.005881 | 0.003105 | 0.003105 |
| 58 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.011518 | 0.011518 | 0.008932 | 0.008932 | 0.005739 | 0.005739 | 0.003089 | 0.003089 |
| 59 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.012223 | 0.012223 | 0.009160 | 0.009160 | 0.005987 | 0.005987 | 0.003270 | nan |
| 60 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.011262 | 0.011262 | 0.008963 | 0.008963 | 0.005760 | 0.005760 | 0.003048 | 0.003048 |
| 61 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.011434 | 0.011434 | 0.008837 | 0.008837 | 0.005587 | 0.005587 | 0.003033 | 0.003033 |
| 62 | 0.012740 | 0.009692 | 0.006063 | 0.003495 | 0.012331 | nan | 0.009331 | nan | 0.006075 | nan | 0.003216 | nan |

*Figure 13: Example of a MDB CSV summary file (snippet showing columns with Rrs matchup variables at 412, 443, 490 and 510nm).*

### 3.2.2 Module RR-AC

#### *3.2.2.1 Overview*

The overall objective of the Atmospheric Correction Round Robin (**RR-AC)** Module is to compare different ocean atmospheric correction methods to support the selection of an atmospheric correction method for the *ocean colour processor*.

Detailed description of OMAPS RR-AC Module is included in the dedicated ATBD.

In general, a reliable and stable atmospheric correction procedure is the basis for ocean colour products of the necessary high quality. According to the OMAPS project methodological requirements, the RR-AC basically follows the method derived for the Ocean Colour CCI and described in detail in [RD-5] and [RD-8].

Figure 14 shows a representation of the RR-AC workflow as implemented in OMAPS. Tables of aggregated data generated with the RR-AC Module can be used as input data, so that aggregation steps can be skipped, and the processing resumes with the combining of data for the statistical analysis and the bootstrapping.

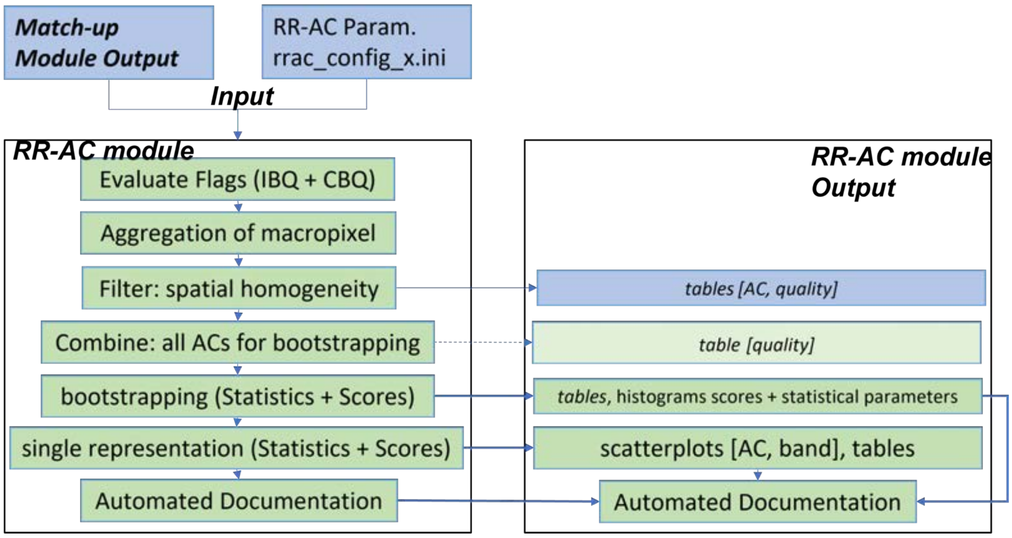



*Figure 14: RR-AC Module of OMAPS project*

The RR-AC requires running the candidates of atmospheric correction processors which shall participate in the RR-AC. Input and output must be prepared individually according to the specification of the particular processor. The AC processor must run on a set of ocean colour Level 1 data to generated marine reflectances (normalised or not normalised). This set of Level 2 products can then be given as input to the MATCHUP Module, which is the start of the RR-AC procedure as described in section 3.2.1.

#### *3.2.2.2 Inside the RR-AC: Aggregation, Filtering, Statistics, scoring and bootstrapping*

Starting from the extracted macropixels of the MATCHUP Module, this data is aggregated into single spectral values per in-situ measurement. During aggregation of the macropixel values the quality flags of the atmospheric correction are applied. A valid pixel expression for each AC must be defined in the parameter file, which can also be combined to a common valid pixel expression (CBQ = common best quality) for all ACs present in the comparison.

Based on the valid pixels, the spatial homogeneity of the macropixel is tested and average values are discarded, if the criterion is not met.

Finally, the RR-AC comprises statistics, scoring and bootstrapping as described in [AD-7], which can be specified in the parameter file.

### *3.2.2.3 Input Data*

#### 3.2.2.3.1 MDB CSV summary files

The input data for this step are the CSV summary files of the MDB products as described in section 3.2.1.4. For each of the Atmospheric Correction candidates in the RoundRobin, a corresponding MDB CSV summary file must be provided as input (to be specified in the configuration file, see below).

#### 3.2.2.3.2 Configuration file

As for the other Offline Processor modules, a configuration file defines a set of parameters which steer the RR-AC processing.

E.g., there is an option to ingest aggregated matchup data, which has been aggregated with the RR-AC Module previously, instead of the macropixel extractions from the MATCHUP Module. This will reduce the processing time, as the aggregation step is skipped.

If valid pixel expressions are changed, the aggregation must be run again.

Although the filtering of spatial homogeneity is a requirement from the matchup protocol (RD36), it is possible to switch it off (set parameter 'check_homogen' to False). This might be necessary, if noisy AC results lead to a loss of almost all data points by this filter (often for a specific band) and no meaningful spectral statistics can be derived.

As for the other modules, the configuration parameters / user options for the RR-AC are provided as configuration file (ini format). This configuration file is again most conveniently generated by the Configuration GUI which is described in more detail in the SUM [AD-5]. Table 3-20 lists the available user options and their meaning, and Figure 15 shows an example for a configuration file used with a RR-AC run .

```
[RR_AC]
; configuration for RoundRobin AC
mdb_input_dir = path/to/mdb_extracts
mdb_input_files_ipf = MDB_S3A_OLCI_L2_OCDB_PML_L2_RRS_IPF_FULL_TEST.csv
mdb_input_files_l2gen = MDB_S3A_OLCI_L2_OCDB_PML_L2_RRS_L2GEN.csv
mdb_input_files_polymer = MDB_S3A_OLCI_L2_OCDB_PML_L2_POLY_TEST.csv
mdb_input_files_sacso = MDB_S3A_OLCI_L2_OCDB_PML_L2_SACSO_TEST.csv
rrac_output_dir_root = path/to/output_directory
processing_label = RRAC_4ACs_test
flag_type = IBQ
rrac_rho_type = rrs
insitu_rho_type = rrs
satellite_rho_type = rrs
check_homogen = False
write_bootstrap_data = True
n_bootstrap = 100
statistical_parameters = MAD,MD,MAPD,MPD,SAM,CHI2
chi2_insitu_bands = rrs_412,rrs_443,rrs_490,rrs_560.5796,rrs_665
chi2_insitu_band_norm = rrs_560.5796
score_parameters = MAD,MD,MAPD,MPD,SAM,CHI2
read_aggregated_data = False
polymer_bits_exclude = bitmask.LAND,bitmask.CLOUD_BASE,bitmask.L1_INVALID,bitmask.NEGATIVE_BB,bitma
polymer_bits_include =
sacso_bits_exclude = flags.BAD
sacso_bits_include =
ipf_bits_exclude = WQSF.CLOUD,WQSF.CLOUD_AMBIGUOUS,WQSF.CLOUD_MARGIN,WQSF.INVALID,WQSF.COSMETIC,WQS
ipf_bits_include = WQSF.WATER
l2gen_bits_exclude = l2_flags.ATMFAIL,l2_flags.LAND,l2_flags.CLDICE,l2_flags.SEAICE
l2gen_bits_include =
```

*Figure 15 Example for a configuration file used with a RR-AC run*

*Table 3-20: Configuration parameters used in the OMAPS RR-AC*

| PARAMETER | DESCRIPTION | POSSIBLE VALUES | DEFAULT |
|---|---|---|---|
| ***mdb_input_dir*** | Directory with MDB CSV summary input files | /path/to/directory | |
| ***mdb_input_files_polymer*** ***mdb_input_files_ipf*** ***mdb_input_files_sacso*** ***mdb_input_files_l2gen*** | CSV matchup data extraction file names, comma separated, split by AC type (polymer, sacso, l2gen, ipf) | e.g., filename1, filename2 | |
| ***rrac_output_dir_root*** | Root output directory for tables, figures and documentation | /path/to/directory | |
| ***processing_label*** | Name of output folder created at *rrac_output_dir_root*; this string is added to all output file names | Any string for a directory name e.g., RRAC_TEST_2 | |
| ***chi2_insitu_bands*** | Names of insitu reflectance bands, which are used in the chi2 analysis, comma separated | e.g., rrs_412, rrs_443, rrs_490, rrs_560.5796, rrs_665 | |
| ***chi2_insitu_band_norm*** | Name of band used to normalise each spectrum | e.g., rrs_560.5796 | |
| ***flag_type*** | Type of quality flags used during aggregation. | IBQ, CBQ | |
| ***insitu_rho_type*** | Type of reflectance of the in-situ data | rrs, rhow | |
| ***satellite_rho_type*** | Type of reflectance of the satellite data | rrs, rhow | |
| ***rrac_rho_type*** | The comparison is based on this type of reflectance. Insitu and satellite data are converted, if necessary. | rrs, rhow | |
| ***check_homogen*** | Boolean, is spatial homogeneity of the macropixel checked according to RD-36. | True, False | True |
| ***write_bootstrap_data*** | Write bootstrap data | True, False | False |
| ***n_bootstrap*** | Number of bootstrap resamplings to analyse distribution of statistics and scores | Any integer. For well-defined distributions N_bootstrap > 10000 | 100 |
| ***statistical_parameters*** | Name of statistical parameter, which is evaluated per band (or on a spectrum) | MAD, MD, MAPD, MPD, SAM, CHI2, r, RMSE.abs, RMSE.rel, bias, residual.error.abs, residual.error.rel, Slope, Intercept, N | MAD,MD, MAPD, MPD,SAM, CHI2 |
| ***score_parameters*** | Statistical parameters, which are analysed and converted into scores | See statistical_parameters | MAD,MD, MAPD,MPD, CHI2 |
| ***read_aggregated_data*** | Read in aggregated data, which has been processed by the RR-AC. Needs also the paths to the original extraction data from the MATCHUP Module (for meta data)! | True, False | False |

| | | | |
|---|---|---|---|
| ***aggregated_data_input_dir*** | Path to folder with aggregated data. | /path/to/directory | |
| ***aggregated_data_input_files*** | Filenames of aggregated data (comma separated) | | |
| ***polymer_bits_exclude*** | Valid pixel expression for AC POLYMER is combined from flag band values. All bands listed in exclude are combined by "AND NOT" | (default)<br>bitmask.LAND, bitmask.CLOUD_BASE, bitmask.L1_INVALID, bitmask.NEGATIVE_BB, bitmask.OUT_OF_BOUNDS, bitmask.EXCEPTION, bitmask.THICK_AEROSOL, bitmask.HIGH_AIR_MASS, bitmask.EXTERNAL_MASK, bitmask.CASE2, bitmask.INCONSISTENCY | |
| ***polymer_bits_include*** | Part of the valid pixel expression for POLYMER data. Combined with "AND". Flag has to be raised in a valid pixel. | | |
| ***sacso_bits_exclude*** | Sacso valid pixel expression. | flags.BAD | |
| ***sacso_bits_include*** | Sacso valid pixel expression. | | |
| ***ipf_bits_exclude*** | IPF valid pixel expression | (default) WQSF.CLOUD, WQSF.CLOUD_AMBIGUOUS, WQSF.CLOUD_MARGIN, WQSF.INVALID, WQSF.COSMETIC, WQSF.SATURATED, WQSF.SUSPECT, WQSF.HISOLZEN, WQSF.HIGHGLINT, WQSF.SNOW_ICE, WQSF.AC_FAIL, WQSF.WHITECAPS, WQSF.ADJAC, WQSF.OC4ME_FAIL, WQSF.RWNEG_O2, WQSF.RWNEG_O3, WQSF.RWNEG_O4, WQSF.RWNEG_O5, WQSF.RWNEG_O6, WQSF.RWNEG_O7, WQSF.RWNEG_O8 | |
| ***ipf_bits_include*** | IPF valid pixel expression | WQSF.WATER | |
| ***l2gen_bits_exclude*** | L2gen valid pixel expression | l2_flags.ATMFAIL, l2_flags.LAND, l2_flags.CLDICE, l2_flags.SEAICE | |
| ***l2gen_bits_include*** | L2gen valid pixel expression | | |

```
[RR_AC]
; configuration for RoundRobin AC
mdb_input_dir = path/to/mdb_extracts
mdb_input_files_ipf = MDB_S3A_OLCI_L2_OCDB_PML_L2_RRS_IPF_FULL_TEST.csv
mdb_input_files_l2gen = MDB_S3A_OLCI_L2_OCDB_PML_L2_RRS_L2GEN.csv
mdb_input_files_polymer = MDB_S3A_OLCI_L2_OCDB_PML_L2_POLY_TEST.csv
mdb_input_files_sacso = MDB_S3A_OLCI_L2_OCDB_PML_L2_SACSO_TEST.csv
rrac_output_dir_root = path/to/output_directory
processing_label = RRAC_4ACs_test
flag_type = IBQ
rrac_rho_type = rrs
insitu_rho_type = rrs
satellite_rho_type = rrs
check_homogen = False
write_bootstrap_data = True
n_bootstrap = 100
statistical_parameters = MAD,MD,MAPD,MPD,SAM,CHI2
chi2_insitu_bands = rrs_412,rrs_443,rrs_490,rrs_560.5796,rrs_665
chi2_insitu_band_norm = rrs_560.5796
score_parameters = MAD,MD,MAPD,MPD,SAM,CHI2
read_aggregated_data = False
polymer_bits_exclude = bitmask.LAND,bitmask.CLOUD_BASE,bitmask.L1_INVALID,bitmask.NEGATIVE_BB,bitma
polymer_bits_include =
sacso_bits_exclude = flags.BAD
sacso_bits_include =
ipf_bits_exclude = WQSF.CLOUD,WQSF.CLOUD_AMBIGUOUS,WQSF.CLOUD_MARGIN,WQSF.INVALID,WQSF.COSMETIC,WQS
ipf_bits_include = WQSF.WATER
l2gen_bits_exclude = l2_flags.ATMFAIL,l2_flags.LAND,l2_flags.CLDICE,l2_flags.SEAICE
l2gen_bits_include =
```

*Figure 16: Example of a configuration file for the OMAPS RR-AC, processing matchup data extractions.*

```
[RR_AC]
; configuration for rr_ac
mdb_input_dir = .
mdb_input_files_polymer =
mdb_input_files_ipf =
mdb_input_files_sacso =
mdb_input_files_l2gen =
rrac_output_dir_root = path/to/output-directory
processing_label = RR_AC_agg
chi2_insitu_bands = rrs_412,rrs_443,rrs_490,rrs_560.5796,rrs_665
chi2_insitu_band_norm = rs_560.5796
flag_type = IBQ
insitu_rho_type = rrs
satellite_rho_type = rrs
rrac_rho_type = rrs
check_homogen = False
write_bootstrap_data = False
n_bootstrap = 100
statistical_parameters = MAD,MD,MAPD,MPD,SAM,CHI2
score_parameters = MAD,MD,MAPD,MPD,CHI2
read_aggregated_data = True
aggregated_data_input_dir = path/to/aggregated-data-folder
aggregated_data_input_files = aggrData_sacso_1.0_IBQ_allCorr_RRAC_4ACs_test.txt,aggrData_ip
```

*Figure 17: Example of a configuration file for the OMAPS RR-AC Module if aggregated data is used as input.*

#### 3.2.2.4 Output Data

The output data from this step consists of:

- Tables of aggregated data (applied: valid pixel expression + spatial homogeneity filter + outlier filter)
- Bootstrap scores (tables and plots)
- Automated documentation

The tables of aggregated data include all the meta data which is needed to ingest the aggregated data on its own and run the RR-AC directly on previously calculated values. The meta data states: the AC processor type and its version (*AC_type, AC_processor_version*), the full path to the original matchup data file (*fullPath_Matchup*), the names of insitu variables and corresponding satellite variables (*insitu_variables, sat_variables*), the type of flagging (*FlagType*), which was applied, and the applied valid pixel expression (*ValidPixelExpression_include, ValidPixelExpression_exclude*) plus whether or not the spatial homogeneity of the macropixel has been checked (*checkHomogeneity*).

```
#AC_type=ipf
#AC_processor_version=06.11
#fullPath_Matchup=D:\Documents\projects\OMAPS\data_email_20210913\\MDB_S3A_OLCI_L2_OCDB_PML_L2_RRS_IPF_FULL_TEST.csv
#insitu_variables=insitu_rrs_412,insitu_rrs_443,insitu_rrs_490,insitu_rrs_510,insitu_rrs_560.5796,insitu_rrs_620.626,ins
#sat_variables=IPF-OL-2_v06.11_Oa02_Rrs,IPF-OL-2_v06.11_Oa03_Rrs,IPF-OL-2_v06.11_Oa04_Rrs,IPF-OL-2_v06.11_Oa05_Rrs,IPF-O
#FlagType=IBQ
#ValidPixelExpression_include=WQSF.WATER
#ValidPixelExpression_exclude=WQSF.CLOUD,WQSF.CLOUD_AMBIGUOUS,WQSF.CLOUD_MARGIN,WQSF.INVALID,WQSF.COSMETIC,WQSF.SATURATE
#checkHomogeneity=False
pixel_ID    central_latitude    central_longitude   central_time    IPF-OL-2_v06.11_latitude    IPF-OL-2_v06.11_longitud
20170216T201814_0_0_0-60f6e4bd1d2b6d0001652599_2315 20.8179 -157.1991   2017-02-16T20:20:07 20.8239 -157.2021   20.8172
20170216T201814_1_0_0-60f6e4bd1d2b6d0001652599_2316 20.8179 -157.1991   2017-02-16T20:20:07 20.8239 -157.2021   20.8172
20170311T202301_0_0_0-60f6e4bd1d2b6d0001652599_2326 20.8193 -157.1791   2017-03-11T20:23:48 20.8251 -157.1816   20.8207
20170315T201845_0_0_0-60f6e4bd1d2b6d0001652599_2328 20.8188 -157.1804   2017-03-15T20:20:06 20.8247 -157.1833   20.8202
20170315T201845_1_0_0-60f6e4bd1d2b6d0001652599_2329 20.8188 -157.1804   2017-03-15T20:20:06 20.8247 -157.1833   20.8202
20170319T201431_0_0_0-60f6e4bd1d2b6d0001652599_2337 20.8216 -157.1996   2017-03-19T20:16:25 20.8276 -157.2032   20.8217
20170322T203835_0_0_0-60f6e4bd1d2b6d0001652599_2339 20.8168 -157.1913   2017-03-22T20:38:35 20.823  -157.1952   20.8172
20170326T203421_0_0_0-60f6e4bd1d2b6d0001652599_2342 20.816  -157.1898   2017-03-26T20:34:54 20.8223 -157.194    20.8172
```

*Figure 18 Excerpt from aggregated data file with meta data header.*

The example of RR-AC results is based on the AC POLYMER (version 4.17) and SASCO (version 1.0). The homogeneity check had to be turned off, otherwise almost all data for rrs at 665nm would be excluded and spectral statistics like the chi square value or the spectral angle mapper (SAM) had to omit the 665nm data.

The number of resamplings for bootstrapping is set to 100 repetitions, for demonstration purposes only. It is recommended to use larger numbers in order to get a clearer analysis of the distributions of statistical parameters and scores.

| Processor | varname | MAD | MD | MAPD | MPD | N |
|---|---|---|---|---|---|---|
| polymer | rrs_412 | 8.24e-04 | 5.42e-05 | 6.2 | -0.5 | 102 |
| polymer | rrs_443 | 5.70e-04 | -2.39e-05 | 5.9 | -0.9 | 102 |
| polymer | rrs_490 | 2.89e-04 | 1.29e-05 | 4.9 | -0.2 | 102 |
| polymer | rrs_510 | 4.08e-04 | 3.9e-04 | 11.6 | 11.1 | 102 |
| polymer | rrs_560.5796 | 7.04e-05 | -2.76e-05 | 5.8 | -2.5 | 102 |
| polymer | rrs_620.626 | 3.83e-05 | -3.64e-05 | 22.8 | -21.8 | 102 |
| polymer | rrs_665 | 1.97e-05 | -4.15e-06 | 22.5 | -6.2 | 102 |
| sasco | rrs_412 | 1.15e-03 | -1.05e-04 | 8.7 | -1.5 | 102 |
| sasco | rrs_443 | 7.56e-04 | -1.56e-04 | 7.9 | -2.2 | 102 |
| sasco | rrs_490 | 7.71e-04 | -7.06e-04 | 13.3 | -12.2 | 102 |
| sasco | rrs_510 | 2.97e-04 | -1.40e-04 | 8.9 | -4.4 | 102 |
| sasco | rrs_560.5796 | 3.30e-04 | -3.21e-04 | 26.6 | -25.8 | 102 |
| sasco | rrs_620.626 | 1.67e-04 | -1.67e-04 | 96.8 | -96.8 | 102 |
| sasco | rrs_665 | 1.65e-04 | -1.65e-04 | 185.9 | -185.9 | 102 |

*Figure 19 Table of statistics per band from the automated documentation (IBQ).*

| Processor | CHI2 | SAM |
|---|---|---|
| polymer | 0.3135 | 0.013 |
| sasco | 0.4801 | 0.0407 |

*Figure 20 Table of spectral statistics from the automated documentation.*

| Processor | rrs_412 | rrs_443 | rrs_490 | rrs_510 | rrs_560.5796 | rrs_620.626 | rrs_665 | CHI2 | SAM | Total.Scores |
|---|---|---|---|---|---|---|---|---|---|---|
| polymer | 2.33 | 2.33 | 4.0 | 0.33 | 4.0 | 4.0 | 4.0 | 2.42 | 3.05 | 26.47 |
| sasco | 1.67 | 1.67 | 0.0 | 3.67 | 0.0 | 0.0 | 0.0 | 1.58 | 0.95 | 9.53 |

*Figure 21 Table of Scores per band or spectral statistics parameter and total scores (from automated documentation).*

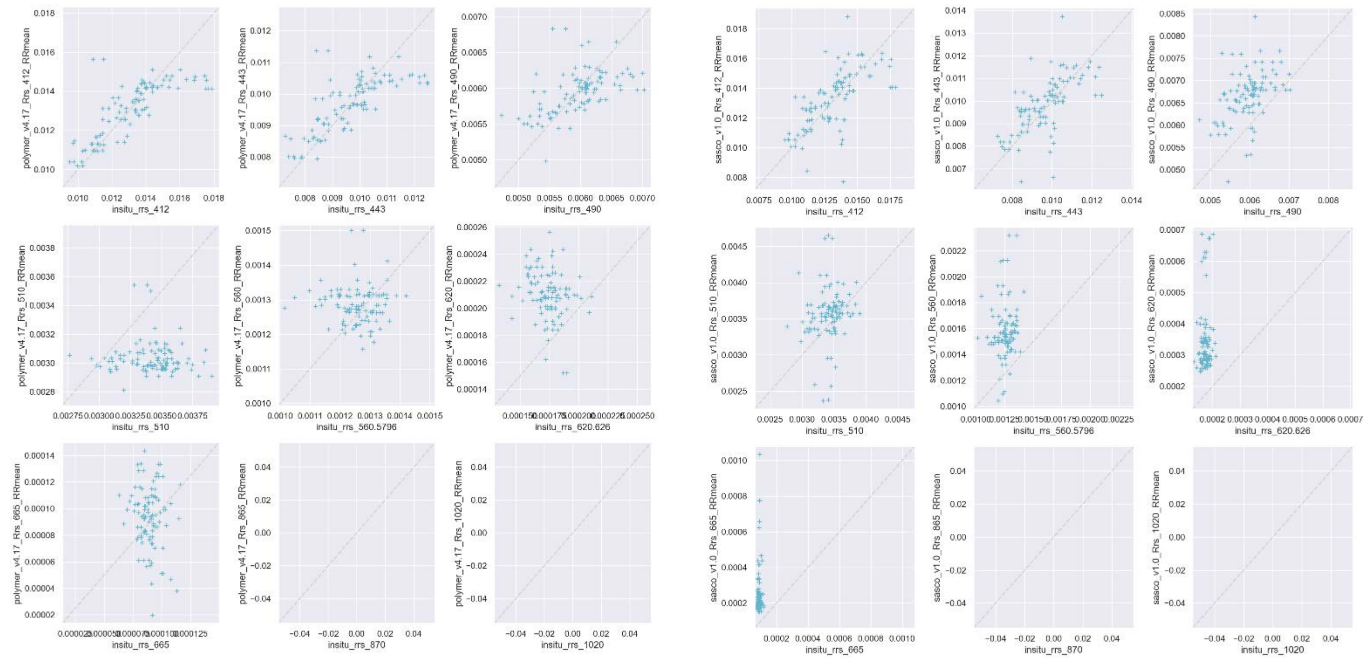


*Figure 22 Scatterplots for POLYMER v4.17 and SASCO v1.0, aggregated data (IBQ, outlier filter, no spatial homogeneity check).*

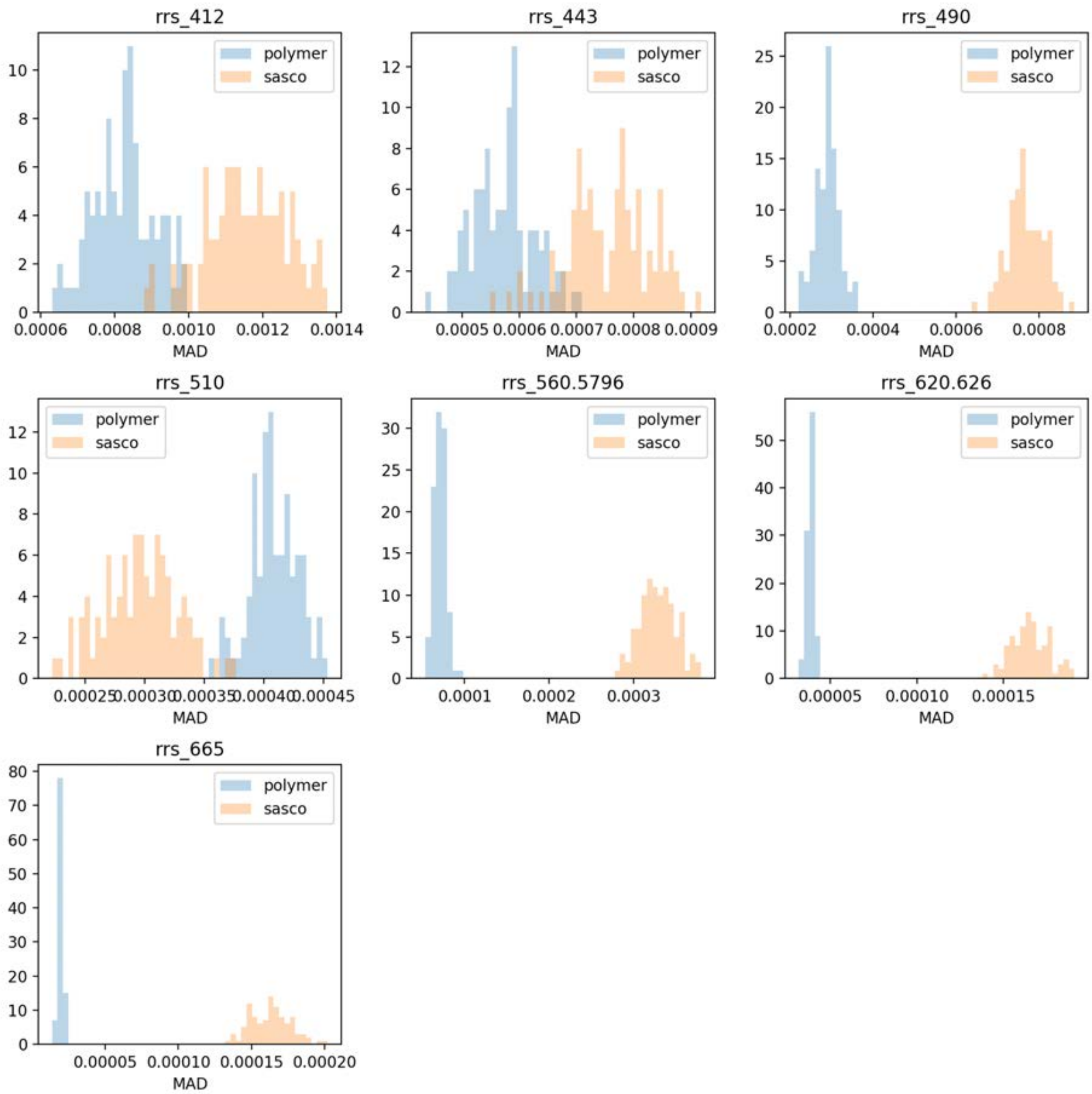



*Figure 23 Example of distributions of MAD (mean absolute difference) derived from bootstrapping (per band).*

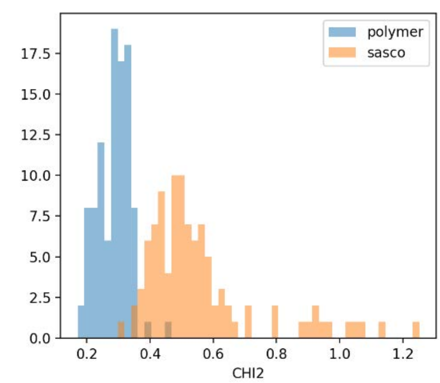



*Figure 24: Distribution of Chi Square Values (from bootstrapping)*

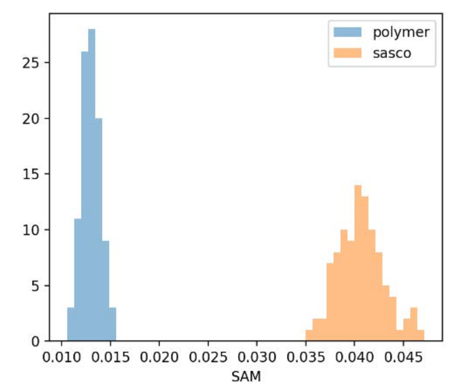



*Figure 25: Distribution of Spectral Angle Mapper (from bootstrapping)*

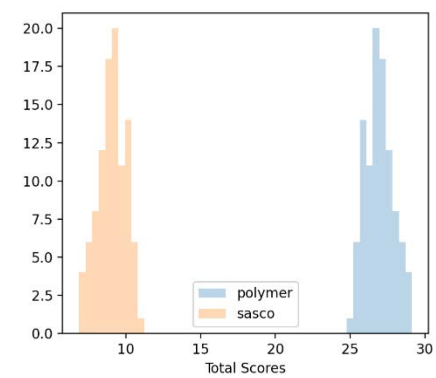



*Figure 26: Distribution of Total Scores*

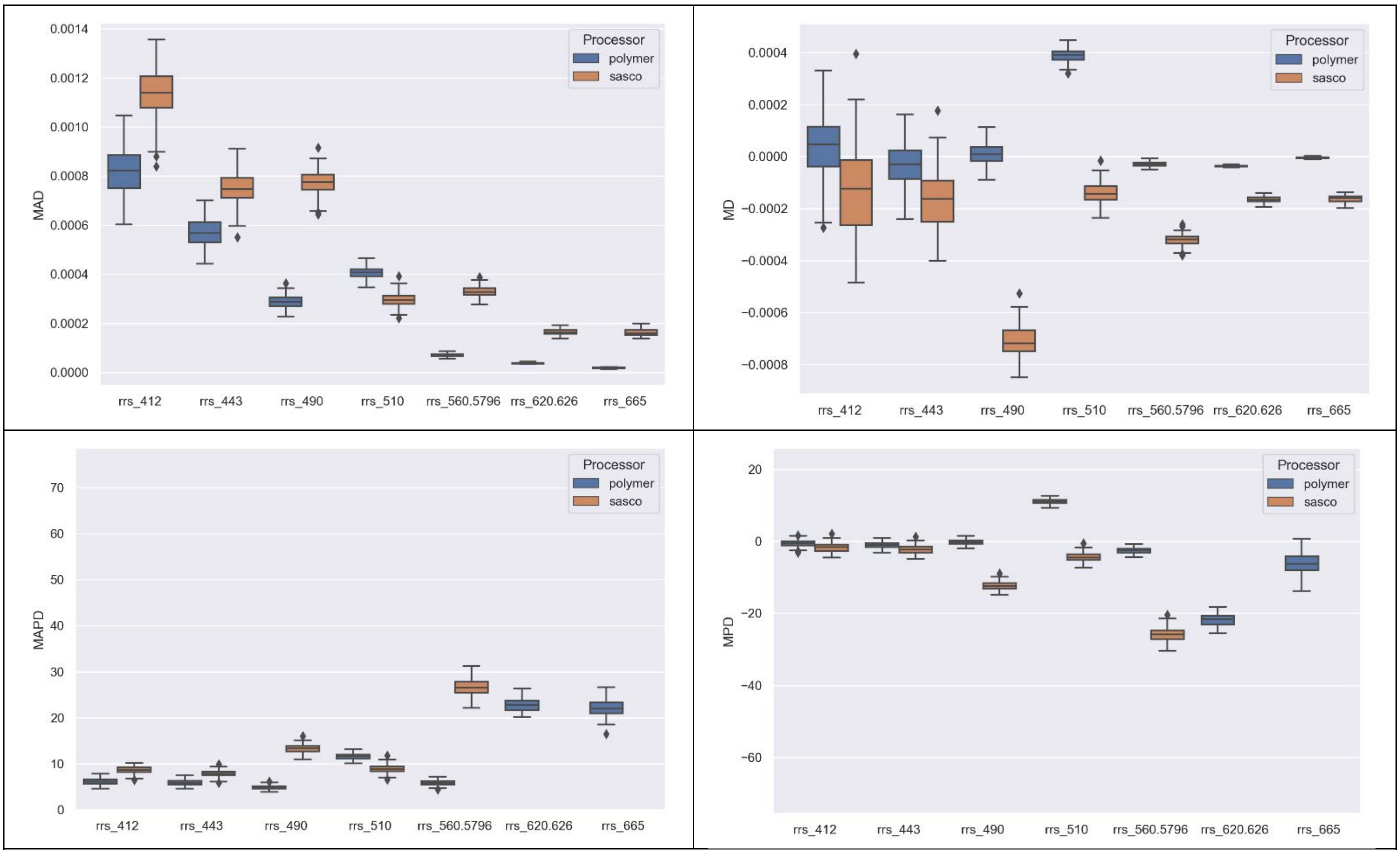



*Figure 27 Distribution of MAD, MD, MAPD, MPD per band and AC processor (from bootstrapping). Some values for SACSO and the bands rrs_620.6 and rrs_665 are missing in MAPD and MP (values up to 200%, see table above)*

## 3.2.3 Module RR-IW

### *3.2.3.1 Overview*

The In-water Round Robin (**RR-IW)** Module has the purpose to objectively compare in-water algorithms which are candidates for the ocean colour processor. The RR-IW is an objective methodology to assess the performance of bio-optical algorithms for use with remotely sensed reflectance (Rrs) ocean colour data. As with the RR-AC, the RR-IW has also been used in OC-CCI. More details are given in [RD-9] and [RD-10].

The RR-IW compares different in-water algorithms. Each algorithm shall take Rrs as input. The Rrs values shall be derived from satellite observations and not from in-situ measurements. The algorithms can be semi-analytical models, designed to retrieve IOPs, and the empirical chlorophyll models.

All algorithms participating in the RR-IW are applied to matchups with reference data. The match-ups generation considers best practices, such as the method by *Bailey* [RD-2] and the match-up protocol developed for Sentinel 3 [RD-6]. The match-up generation as described in section 3.2.1 is used.

The RR-IW Module uses the matchups generated by the MATCHUP Module as input. This requires processing of ocean colour products for the algorithms included in the RR-IW using the ocean processing subsystem and running the MATCHUP Module on these outputs.

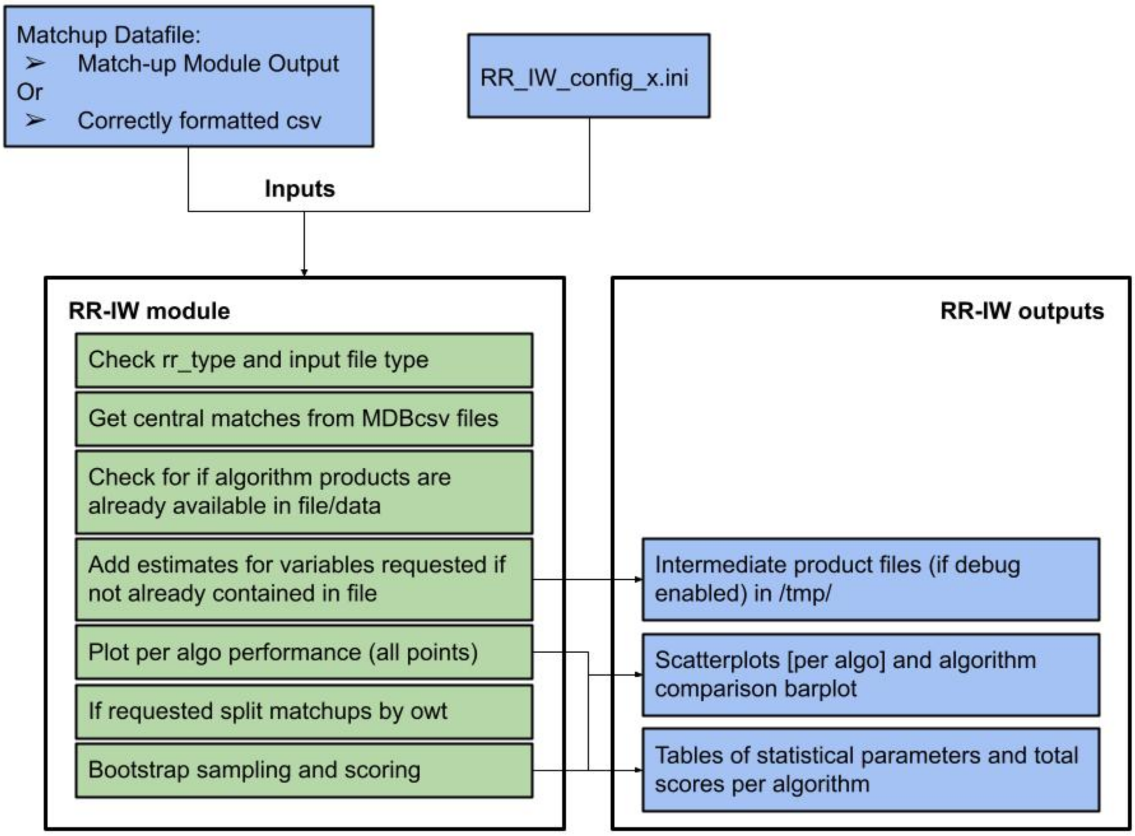



*Figure 28: Overview of the In water round robin comparison procedure with input files, processing stages and output files.*

The RR-IW produces tabulated output (text file) containing the statistics and scores for the whole match-up data set and per optical water type. Plots for the scores as in [RD-10] are also generated.

### 3.2.3.2 Input Data

#### 3.2.3.2.1 MDB CSV summary files

The primary source of input data for the IW_RR Module is CSV summary files of the MDB products as described in section 3.2.1.4. The location of this file is specified in the in-water-round-robin (IWRR) configuration file (see section 3.2.3.2.3) along with the input file type (‘MDBcsv’ in this case). An example of such a file is shown below, suitable for an assessment of ocean colour algorithms producing chlorophyll-a products.

#creation_time=2021-10-08T13:17:29.668397
#satellite=Sentinel-3
#sensor=OLCI
#description=Sentinel-3 OLCI - OCDB Matchups Data Base
#data_policy=tbd
#AC_processor=polymer
#AC_processor_version=4.17
#sat_bands=Rrs_400;Rrs_412;Rrs_443;Rrs_490;Rrs_510;Rrs_560;Rrs_620;Rrs_665;Rrs_681;Rrs_709;Rrs_754;Rrs_779;Rrs_865;Rrs_1020
#macro_pixel_size=5
#sat_variables=Rrs_412;chlor_oc2;chlor_oc3;chlor_oc4;chlor_ocx;Rrs_443;Rrs_490;Rrs_510;Rrs_560;Rrs_620;Rrs_665;chlor_oci2
#insitu_variables=chla_fluor;chla_hplc
#MDAD_of_OC_variables=0.888120;0.048609
#MDAPD_of_OC_variables=99.618316;33.293819

| pixel_ID | central_latit | central_long | central_time | satellite_lati | satellite_lon | insitu_lat | insitu_lon | insitu_time | time_differe | satellite_nur | insitu_chla_f | satellite_Rrs_412 | satellite_Rrs_ | satellite_Rrs | satellite_Rrs_412 | insitu_chla_h | satellite_chl | satellite_chl | satellite_chl | satellite_chl |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.758099 | -158.0069 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | 0.014953 | 0.014844 | 0.000462 | 0.014953 | nan | 0.049224 | 0.047768 | 0.003649 | 0.049224 |
| S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.757466 | -158.00354 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | 0.014404 | 0.014844 | 0.000462 | 0.014404 | nan | 0.049252 | 0.047768 | 0.003649 | 0.049252 |
| S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.756832 | -158.00019 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | 0.014512 | 0.014844 | 0.000462 | 0.014512 | nan | 0.046815 | 0.047768 | 0.003649 | 0.046815 |
| S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.756199 | -157.99685 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | 0.015359 | 0.014844 | 0.000462 | 0.015359 | nan | 0.047342 | 0.047768 | 0.003649 | 0.047342 |
| S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.755566 | -157.9935 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | nan | 0.014844 | 0.000462 | 0.013928 | nan | nan | 0.047768 | 0.003649 | 0.065384 |
| S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.755503 | -158.00746 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | 0.014401 | 0.014844 | 0.000462 | 0.014401 | nan | 0.049057 | 0.047768 | 0.003649 | 0.049057 |
| S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.754869 | -158.00411 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | 0.014173 | 0.014844 | 0.000462 | 0.014173 | nan | 0.050391 | 0.047768 | 0.003649 | 0.050391 |
| S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.754235 | -158.00076 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | 0.014073 | 0.014844 | 0.000462 | 0.014073 | nan | 0.053542 | 0.047768 | 0.003649 | 0.053542 |
| S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.753602 | -157.99742 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | 0.015461 | 0.014844 | 0.000462 | 0.015461 | nan | 0.041306 | 0.047768 | 0.003649 | 0.041306 |
| S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.752962 | -157.99403 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | 0.015361 | 0.014844 | 0.000462 | 0.015361 | nan | 0.054472 | 0.047768 | 0.003649 | 0.054472 |
| S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.752908 | -158.00803 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | nan | 0.014844 | 0.000462 | 0.016094 | nan | nan | 0.047768 | 0.003649 | 0.051552 |
| S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.752274 | -158.00468 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | 0.015333 | 0.014844 | 0.000462 | 0.015333 | nan | 0.045879 | 0.047768 | 0.003649 | 0.045879 |

*Figure 29: Example of a matchup file suitable for chlorophyll-a algorithm round robin*

#### 3.2.3.2.2 CSV files

The IW_RR Module also has the capability to ingest non-MDBcsv files provided that they contain correctly named variables as columns and with a single point matchup between the in-situ variable and the satellite Rrs data per row. In this case the file type specified in the configuration file should 'csv'. An example of such a file is shown below, but we recommend using MDBcsv outputs as the primary source of IW-RR comparisons.

| time | lat | lon | chla_hplc | chla_fluor | etopo1 | long_adjust | Rrs_412 | Rrs_443 | Rrs_490 | Rrs_510 | Rrs_560 | Rrs_665 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1997-09-25 | 31.7826 | -17.8372 | 0.049 | | 4507 | 1997-09-25 | 0.01324 | 0.010705 | 0.006808 | 0.003696 | 0.001548 | 9.26E-05 |
| 1997-09-25 | 31.475 | -18.2116 | 0.033 | | 4636 | 1997-09-25 | 0.012646 | 0.010191 | 0.006705 | 0.003632 | 0.001511 | 0.000103 |
| 1997-09-25 | 31.1666 | -18.606 | 0.037 | | 4684 | 1997-09-25 | 0.013251 | 0.010693 | 0.006911 | 0.00373 | 0.001538 | 9.66E-05 |
| 1997-09-26 | 30.8372 | -18.9853 | 0.037 | | 4708 | 1997-09-25 | 0.012772 | 0.010201 | 0.00647 | 0.003435 | 0.001437 | 9.72E-05 |
| 1997-09-26 | 30.5009 | -19.3523 | 0.038 | | 4758 | 1997-09-26 | 0.014796 | 0.011634 | 0.007359 | 0.004065 | 0.001568 | 0.00013 |
| 1997-09-26 | 20.7928 | -157.207 | 0.20313 | | 1376 | 1997-09-26 | 0.009697 | 0.007893 | 0.005749 | 0.003192 | 0.001359 | 7.39E-05 |

*Figure 30: Example of 'generic' csv file that would be suitable for ingestion by the IW-RR Module.*

#### 3.2.3.2.3 Configuration file

The in-water round robin also uses a configuration file to set options such as the algorithms to assess, how to split the water class assessment etc. An example of such a configuration file is shown below.

```
[Bands]
Band1=412
Band2=443
Band3=490
Band4=510
Band5=560
Band6=665
all_bands=412,443,490,510,560,665
[matchup_files]
match_file_chl=/tmp/OMAPS_IW_RR/matchups_IW_RR_polymer/MDB_S3A_OLCI_L2_OCDB_IW_RR_L2_FULLTEST_CHLA.csv
match_file_iop=NaN
match_file_kd=NaN
match_file_rrs=/tmp/OMAPS_IW_RR/matchups_IW_RR_polymer/MDB_S3A_OLCI_L2_OCDB_IW_RR_L2_FULLTEST_CHLA.csv
# naming convention?
[candidate_algorithms]
chl_algorithms: chlor_oc2,chlor_oc3,chlor_oc4,chlor_ocx,chlor_oci,chlor_oci2
[RR_options]
Sensor: OLCI
RR_type: chl
insitu_column_string: chla_fluor
bootstraps: 100
LogNormVar: True
SplitByWaterClass: True
WeightedByWaterClass: False
threshold_memb: 0.3
Split_Criteria: dominant
[output_options]
Plots: True
Tables: True
Directory: /tmp/OMAPS_IW_RR/output_split_by_wc/
Debug: False
```

*Figure 31: Example configuration file for the IW-RR chlorophyll-a algorithm intercomparison*

*Table 3-21: Main properties of RR-IW input data set*

| | |
|---|---|
| ***Processing step*** | **In Water Round Robin** |
| ***Status in processing*** | Input data |
| ***Naming conventions*** | ocdb_extraction_[LABEL]_[OCDB_DATASET_ID].csv<br>e.g.<br>'ocdb_extraction_FIRSTTEST_5ece9a501d2b6d0001666a26.csv' |
| ***Storage*** | Local |
| ***Format*** | CSV File |
| ***Temporal Coverage*** | 1997-01-02T12:48:00 - 2018-04-26T20:00:00 |
| ***Spatial Coverage*** | latitude: [-90.0, 90.0], longitude: [-180.0, 180.0] |
| ***Dimensions*** | single data points (lat, lon, time) |
| ***Number of data points in file*** | >60000 |
| ***Number of parameters per data point*** | >10 |
| ***Size per file*** | ~30MB |

*Table 3-22: Parameters in RR-IW input data set (not all variables are required as in-situ measurements are only required for the corresponding OC-products to be assessed).*

| Parameter | Unit | Data Type | Size (Bytes) | Description |
|---|---|---|---|---|
| Pixel_ID | | String | 68 | Identifier combining the Granule string, the pixel location within the micropixel and the matchup database file for the extraction.<br>Example:<br>S3A_OL_1_EFR____20170523T202832_0_0_0-60f14db91d2b6d0001652594_1975 |
| insitu_<variable> | variable | double | 8 | matching in situ variable as specified in the configuration |
| satellite_<variable> | variable | double | 8 | matching satellite variable as specified in the configuration |
| satellite_num_valid_pixels_<reference variable> | none | int | 2 | number of valid pixels in the micropixel for the 'reference' matchup variable (the variable on which the filter criteria are applied) specified in the configuration |
| satellite_<flags> | none | byte or int | 1/2 | satellite flag variables (depending on OC product type and algorithm). E.g., 'satellite_bitmask' in case of Polymer products. |
| water_class<n> (optional) | none | double | 8 | water classes obtained from satellite (depending on OC product type and algorithm). E.g., n=1,..,14 for Polymer products. Optional output (user defined). |

| PARAMETER | UNIT | DATA TYPE | SIZE (BYTES ) | DESCRIPTION |
|---|---|---|---|---|
| satellite_latitude | deg | doubl e | 8 | satellite per-pixel latitude |
| satellite_longitude | deg | doubl e | 8 | satellite per-pixel longitude |
| central_latitude | degrees_nort h | float | 8 | macropixel central latitude |
| central_longitude | degrees_east | float | 8 | macropixel central longitude |

### 3.2.3.3 *Output Data*

- tabulated output (text files)
- Score plots
- Duplicate of configuration used for RR exercise in output directory

The candidate in water algorithms are applied to the matched Rrs and the outputs provided are 1) If Debug option is enabled than a copy of the input matchup file with the algorithm results and water class memberships appended is saved, 2) A summary table of the Round Robin scores per algorithm and 3) summary plots showing the algorithm performance. Examples are given below:

Output data with per algorithm chlorophyll-a estimates and water class memberships.

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | pixel_ID | central_latit | central_long | central_time | latitude | longitude | insitu_lat | insitu_lon | insitu_time | time_differe | num_valid_p | insitu_chla_f | Rrs_412 | Rrs_412_me | Rrs_412_std | Rrs_412_ori | insitu_chla_t | chlor_oc2 | chlor_oc2_m | chlor_oc2_st | chlor_oc2_or | b |
| 2 | S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.751641 | -158.00133 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | 0.015214 | 0.014844 | 0.000462 | 0.015214 | | 0.049189 | 0.047768 | 0.003649 | 0.049189 | |
| 3 | S3A_OL_1_E | 22.7516 | -158.0013 | 2017-05-23T | 22.751641 | -158.00133 | 22.75 | -158 | 2017-05-23T | 23671 | 22 | 0.07 | 0.01504 | 0.01479 | 0.000399 | 0.01504 | | 0.046941 | 0.046064 | 0.004566 | 0.046941 | |
| 4 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.9907 | -122.2838 | 2017-06-04T | 41400 | 23 | 0.73564 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 5 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.8845 | -122.2178 | 2017-06-04T | 38249 | 23 | 0.706382 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 6 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.7795 | -122.1525 | 2017-06-04T | 35239 | 23 | 0.618607 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 7 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.67 | -122.0844 | 2017-06-04T | 32111 | 23 | 0.708358 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 8 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.5696 | -122.0219 | 2017-06-04T | 29112 | 23 | 1.170337 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 9 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.4608 | -121.9547 | 2017-06-04T | 25969 | 23 | 0.582661 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 10 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.4069 | -121.9215 | 2017-06-04T | 24406 | 23 | 0.457267 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 11 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.3521 | -121.8876 | 2017-06-04T | 22839 | 23 | 0.566778 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 12 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.2318 | -121.8132 | 2017-06-04T | 19347 | 23 | 0.844315 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 13 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.1202 | -121.7443 | 2017-06-04T | 16114 | 23 | 1.613393 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 14 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.018 | -121.6816 | 2017-06-04T | 13050 | 23 | 2.023011 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 15 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.8989 | -121.6079 | 2017-06-04T | 9398 | 23 | 6.261303 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 16 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.7824 | -121.5364 | 2017-06-04T | 5806 | 23 | 3.778517 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 17 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.7248 | -121.5011 | 2017-06-04T | 4067 | 23 | 0.744 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 18 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.5487 | -121.3932 | 2017-06-04T | 1182 | 23 | 0.919551 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 19 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.4347 | -121.3237 | 2017-06-04T | 4712 | 23 | 1.471281 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 20 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.374 | -121.2868 | 2017-06-04T | 6613 | 23 | 1.404404 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 21 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.3196 | -121.2533 | 2017-06-04T | 8263 | 23 | 0.940449 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 22 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.2002 | -121.1804 | 2017-06-04T | 12038 | 23 | 0.681303 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 23 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.1201 | -121.1318 | 2017-06-04T | 14539 | 23 | 0.656225 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 24 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 33.9952 | -121.4263 | 2017-06-05T | 22700 | 23 | 3.143191 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 25 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.0375 | -121.4523 | 2017-06-05T | 24018 | 23 | 2.303056 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 26 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.2183 | -121.5636 | 2017-06-05T | 29466 | 23 | 1.191236 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 27 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.2722 | -121.5968 | 2017-06-05T | 31098 | 23 | 1.488 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 28 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.3874 | -121.6676 | 2017-06-05T | 34809 | 23 | 2.269618 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 29 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.5053 | -121.7401 | 2017-06-05T | 38832 | 23 | 0.706382 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 30 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.5627 | -121.7756 | 2017-06-05T | 40789 | 23 | 0.580989 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 31 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 34.6202 | -121.811 | 2017-06-05T | 42787 | 23 | 0.873573 | 0.003133 | 0.003621 | 0.000401 | 0.003133 | | 0.633139 | 0.597793 | 0.038545 | 0.633139 | |
| 32 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.9907 | -122.2838 | 2017-06-04T | 41399 | 23 | 0.73564 | 0.003037 | 0.003479 | 0.000433 | 0.003037 | | 0.631669 | 0.59818 | 0.050495 | 0.631669 | |
| 33 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.8845 | -122.2178 | 2017-06-04T | 38248 | 23 | 0.706382 | 0.003037 | 0.003479 | 0.000433 | 0.003037 | | 0.631669 | 0.59818 | 0.050495 | 0.631669 | |
| 34 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.7795 | -122.1525 | 2017-06-04T | 35238 | 23 | 0.618607 | 0.003037 | 0.003479 | 0.000433 | 0.003037 | | 0.631669 | 0.59818 | 0.050495 | 0.631669 | |
| 35 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.67 | -122.0844 | 2017-06-04T | 32110 | 23 | 0.708358 | 0.003037 | 0.003479 | 0.000433 | 0.003037 | | 0.631669 | 0.59818 | 0.050495 | 0.631669 | |
| 36 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.5696 | -122.0219 | 2017-06-04T | 29111 | 23 | 1.170337 | 0.003037 | 0.003479 | 0.000433 | 0.003037 | | 0.631669 | 0.59818 | 0.050495 | 0.631669 | |
| 37 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.4608 | -121.9547 | 2017-06-04T | 25968 | 23 | 0.582661 | 0.003037 | 0.003479 | 0.000433 | 0.003037 | | 0.631669 | 0.59818 | 0.050495 | 0.631669 | |
| 38 | S3A_OL_1_E | 35.9898 | -122.2845 | 2017-06-04T | 35.989809 | -122.28447 | 35.4069 | -121.9215 | 2017-06-04T | 24405 | 23 | 0.457267 | 0.003037 | 0.003479 | 0.000433 | 0.003037 | | 0.631669 | 0.59818 | 0.050495 | 0.631669 | |

*Figure 32: Example output csv when debug option is enabled.*

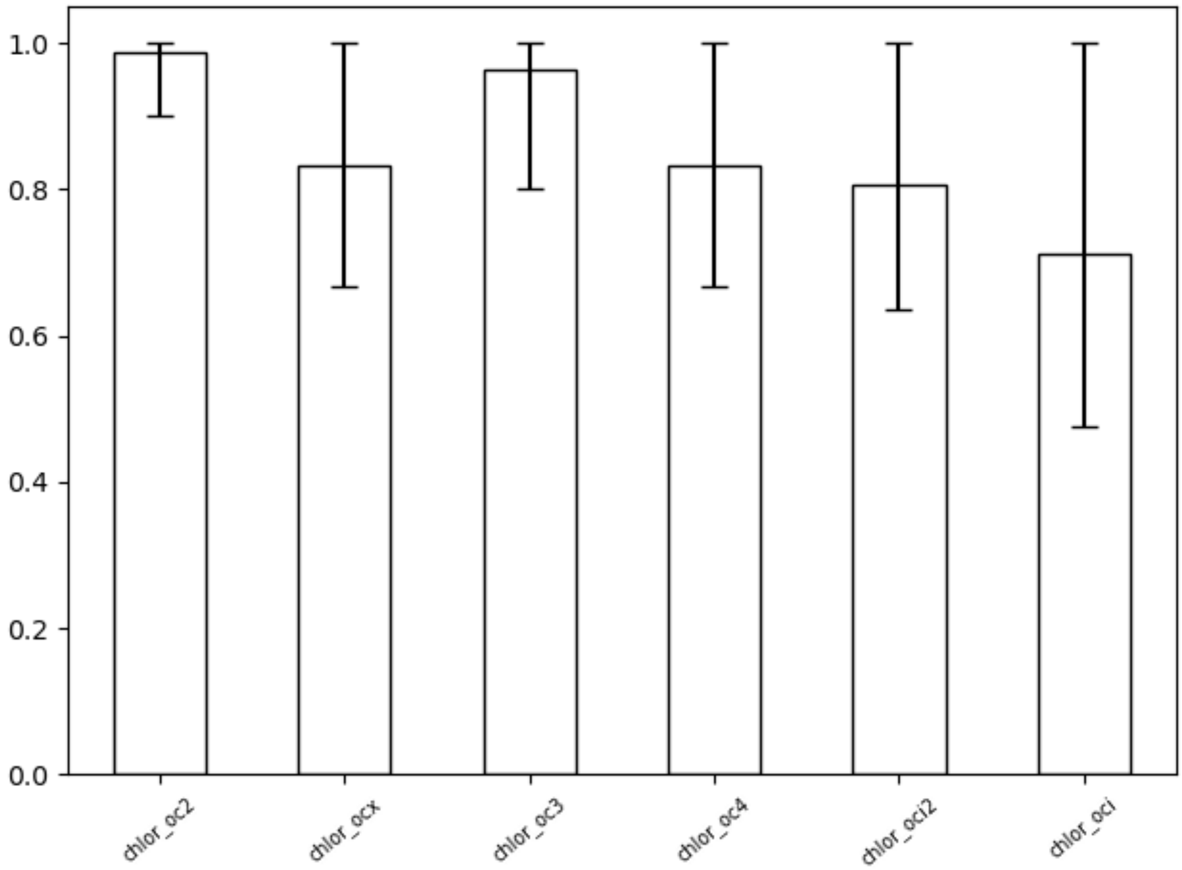
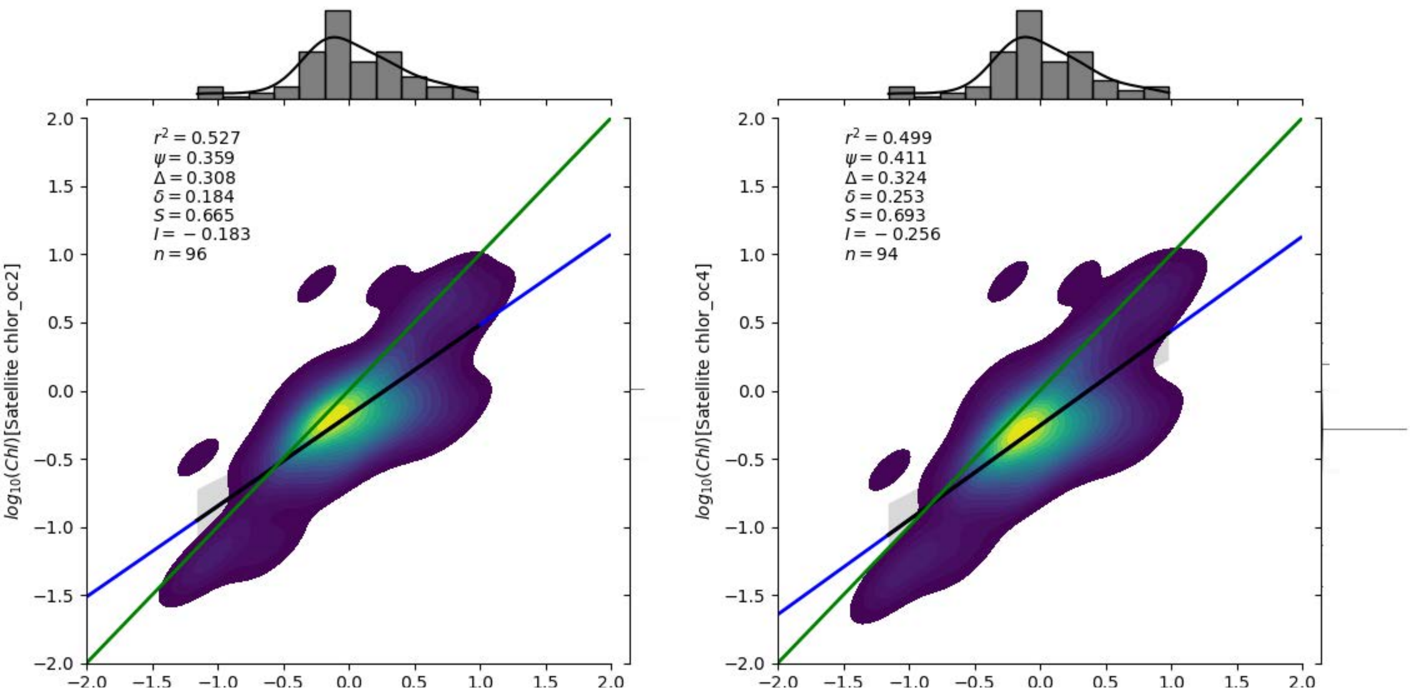



*Figure 33: Example summary plot of multi-metric performance scores across a range of algorithms with bootstrapping.*



### 3.2.4 Module SVC

#### 3.2.4.1 Overview

System Vicarious Calibration (**SVC**) is a crucial step in the processing of ocean colour missions to meet the stringent requirements in absolute accuracy and across-sensor consistency of the Ocean Colour Radiometry. The methodology consists in using high-quality sea truth radiometric measurements comparable to satellite data (match-up process; Req-SVC-001), propagate them at TOA and compute multiplicative gains on the TOA radiometry. SVC gains are computed for all matchups of the calibration dataset and then temporally averaged. They adjust for potential systematic bias in the sensor+processor system, assuming any temporal drift has been corrected beforehand in the Level-1 instrumental calibration.

The SVC Module calculates gains for a given sensor and processor combination, the latter comprising notably the atmospheric correction (AC). In the visible, these SVC gains minimise the bias between the output of the AC and reference (in-situ) measurements, expressed in remote-sensing reflectance (Rrs). The module also computes gains in the NIR, through exactly the same methodology but using only a different dataset and parameterization of the AC (oligotrophic waters, assumption on aerosol model, Req-SVC-002). A more comprehensive summary of the methodological requirements for OMAPS and the underlying method is given in [AD-4].

The SVC Module is strictly identical to the OC-SVC-TOOL software used by EUMETSAT (https://www.eumetsat.int/ocean-colour-system-vicarious-calibration-tool). This software has been designed in a generic way to handle possibly any Level-2 processor. Any processor is considered as a black box (Figure 34) with:

- Level-1b products as input;
- ADFs as parameterization of the processing (here specifically the ADF related to SVC gains);
- Level-2 product as output.

The interface between the generic module and any specific Level-2 processor is ensured by a so-called Level-2 wrapper. The wrapper is an executable file (e.g., script shell, python, etc.) which launches the Level-2 processor for given inputs data, possibly applies some post-processing not included in the Level-2 processor (e.g., BRDF normalization) and provides Level-2 extraction in the standard MDB format defined by EUMETSAT (see below). Hence the Level-2 wrapper does not strictly belong to the SVC Module and must be provided together with the specific Level-2 black-box. More description on the Level-2 wrapper can be found in the SVC-tool user guide [RD-11].

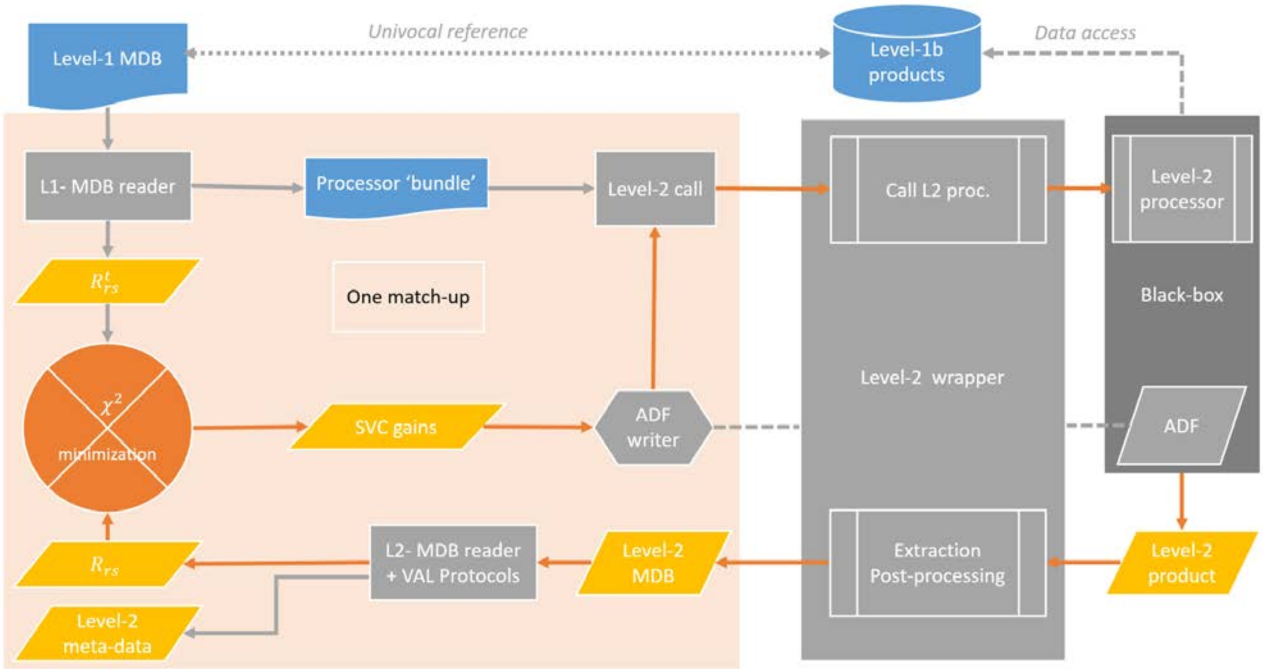



*Figure 34 Schematic of the SVC gains computation. left: the SVC module for individual gain (one match-up); right: the Level-2 processor and associated wrapper interfacing with the SVC module.*

Full details about the OC-SVC-TOOL are given in this user manual [RD-11], which shall be the reference for the OMAPS SVC Module: input/output data, installation, operations manual. Sections below only summarises the main input/output data requirements.

### 3.2.4.2 Input Data

In a standard use, the SVC Module is launched through a graphical user interface (GUI), allowing selection of the match-up database, directory where Level-1 PDUs are stored, Level-2 processor (through the wrapper) and options of the SVC run (Figure 35). All fields of the GUI refer to input data described below. Alternatively, the module can be launched in batch mode with same fields as parameters.

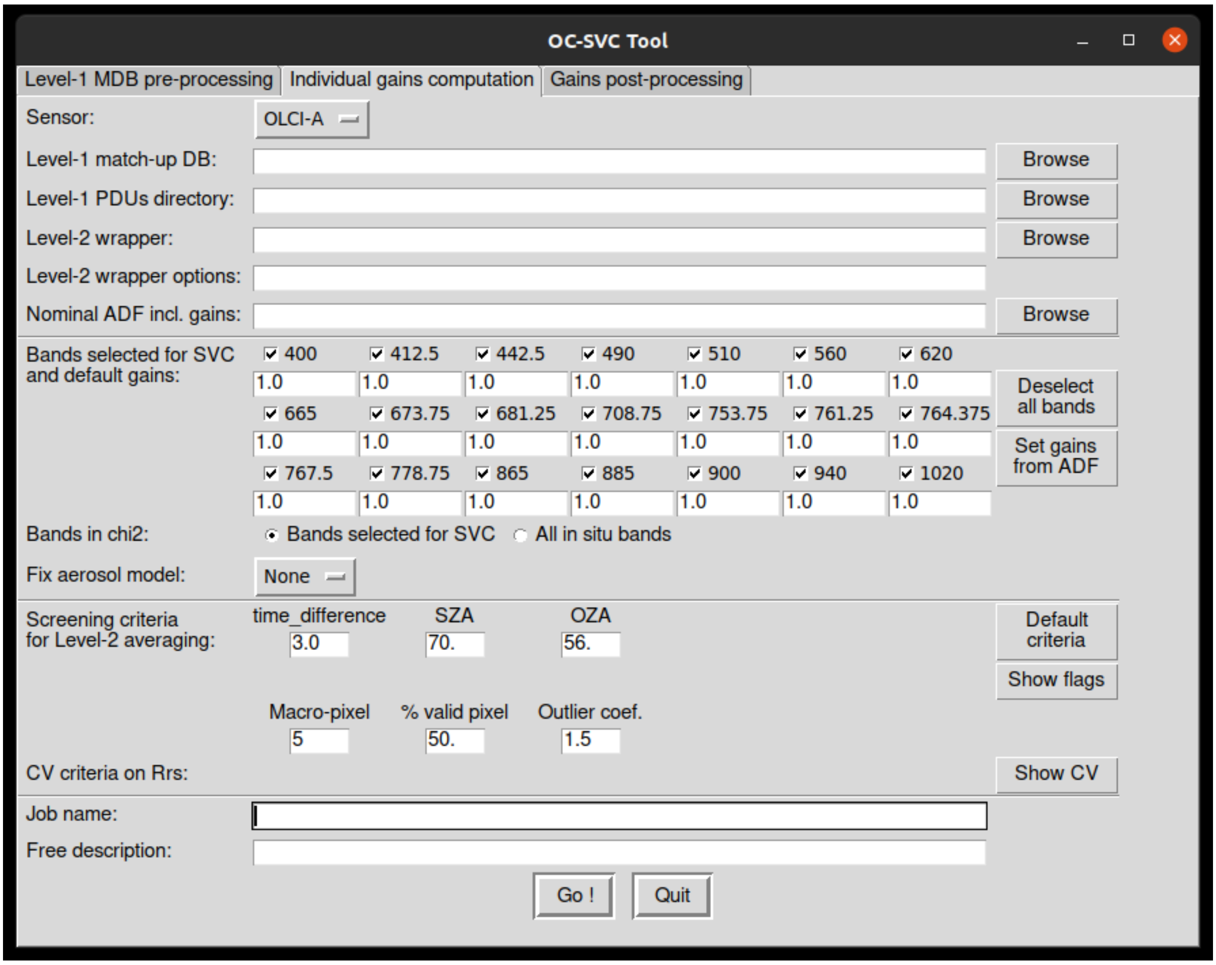



*Figure 35 GUI of the SVC Module, identifying the input of a SVC job*

- Level-1 Match-ups

The Level-1 match-up dataset follows the EUMETSAT match-up database format (MDB). It contains notably the in-situ data and concurrent satellite data. Two processing modes are available, with different requirements for the Level-1 MDB:

- In PDU processing mode, the per-pixel extraction is not used by the SVC Module because it is assumed that the OC processor can only handle native satellite Level-1 format (PDUs), not pixel extractions. In this case, only the PDU name is required as satellite information and the Level-1 match-up database hence essentially defines the pairing between the in-situ data and the corresponding Level-1 PDUs.
- In CSV processing mode, each Level-1 match-ups is converted into temporary CSV files (with semi-colon as delimiter and limited to the macro-pixel chosen by the user). The CSV file is then processed instead of the native PDU by the Level-2 processor. This option can be useful to save CPU time when the processor beneficiates from such capability. In this case, all Level-1 data required by the processor shall be present in the MDB (radiometry, geolocation, ancillary data, flags…). Importantly, this processing mode directly ingest the Level-1 data stored in the MDB and does not extract the native Level-1 PDUs. Hence it is the responsibility of the user to be sure that the data in the Level-1 MDB effectively correspond to the Level-1 PDUs on the disk.

*Table 3-23: Main properties of SVC input: Level-1 match-up database*

| | |
|---|---|
| ***Processing step*** | **SVC** |
| ***Status in processing*** | Input data |
| ***Naming convention*** | MDB_A_L1_$SITE_$START_$STOP.nc |
| ***Storage*** | TBD |
| ***Format*** | netCDF |
| ***Temporal Coverage*** | As long as possible over the mission lifetime (indicated in the MDB filename through START and STOP times) |
| ***Spatial Coverage*** | Calibration site (typically: MOBY site) |
| ***Dimensions*** | OLCI: 25x25 FR pixels |
| ***Size per file*** | Roughly 230KB per match-up (typically less than 30MB for OLCI matchups at MOBY in 3 years) |
| ***Number of files*** | One file per calibration site (typically one for MOBY) |
| ***Processing version*** | Same Level-1 processing baseline than that used for the Level-1 PDUs and for all processing in OMPAS (RR-AC, Validation…) |

*Table 3-24: Parameters in SVC input: Level-1 match-up database when SVC in run in PDU mode (compliant with EUMETSAT MDB format)*

| Parameter | Unit | Data Type | Description |
|---|---|---|---|
| satellite_PDU | none | string | OLCI source PDU name |
| insitu_time | second | double | In situ acquisition time, in seconds since 1970-01-01 |
| time_difference | seconds | double | Absolute time difference between satellite acquisition and in situ acquisition |
| insitu_Oaxx_Rrs | $sr^{-1}$ | double | Above water in situ Remote Sensing Reflectance at band xx matching satellite band |

- Directory of Level-1 PDUs

The directory containing all Level-1 PDUs covering the calibration site shall be provided as input (recursive search applies to sub-directories). The format of the Level-1 PDUs shall be that used by the Ocean Colour Processor: either full granule or minifiles.

- Nominal ADF

The nominal ADF is that containing the vicarious gains and used by the Ocean Colour Processor. A copy is created in the local directory of the SVC Module (output directory), with gains modified in the iterative process until convergence. Importantly, the gains of bands not selected in the SVC process (i.e., bands not calibrated by the SVC Module, see option below) remains as it in the nominal ADF; in particular the ADF containing gains computed in the NIR only shall be used in a second stage to compute gains in the VIS. When the processor does not read the gain in a specific ADF but directly takes a list of arguments as input (e.g., POLYMER with parameter “calib”), it is the role of the wrapper to interface between this mandatory ADF and the processor.

- Level-2 data

Level-2 data are not a direct input per se, but an intermediate data, generated by the Ocean Colour Processor and the Level-2 wrapper, and read by the SVC Module. The Level-2 data should correspond to a match-up extraction of the processed PDU. The Level-2 output must be a file name MDB_L2.nc, formatted into a MDB consistent with EUMETSAT Level-2 MDB in term of format (variables, attributes). The strictly required fields are given below. Other Level-2 outputs can be added (e.g., chlorophyll, aerosol products…) under the same naming convention *satellite_variable*, that will be used for further screening.

*Table 3-25: Required parameters of Level-2 MDB intermediate data (MDB_L2.nc)*

| PARAMETER | UNIT | DATA TYPE | SIZE (BYTES) | DESCRIPTION |
|---|---|---|---|---|
| satellite_Oaxx_Rrs | $sr^{-1}$ | double | 8 | satellite Remote Sensing Reflectance at band xx, corrected for BRDF |
| satellite_WQSF | none | uint | 2 | Classification flags, quality and science flags for Marine and Inland Waters pixels |

- Options of the SVC job

A series of options define the SVC job. The first is the list of bands to be calibrated by the SVC Module during a run. In case of sequential runs (e.g., NIR computation followed by VIS computation), two jobs shall be launched with their respective options. Other fields refer to thresholds and flags to be considered in the averaging process at match-up level. The exact definition of thresholds is given in the job configuration file under following template:

[SVC]

- thresholds = var1, var2, …, varN
- var1 = threshold of var1
- …
- varN = threshold of varN

The GUI automatically adapts to the list of thresholds. Then, if the user provides a value of below or equal to zero, or let the field empty, the criterion is not applied in the screening.

*Table 3-26: Parameters in SVC input: options of the SVC job. Thresholds are only given as example.*

| PARAMETER | UNIT | DATA TYPE | SIZE (BYTES) | DESCRIPTION |
|---|---|---|---|---|
| bands | nm | double | 8 | List of bands to be calibrated (comma separated) |
| time_difference | seconds | double | 8 | Absolute time difference between satellite acquisition and in situ acquisition |
| SZA_max | degree | double | 8 | Maximum Sun zenith angle of the satellite acquisition |
| OZA_max | degree | double | 8 | Maximum observation zenith angle of the satellite acquisition |
| Chl_max | $mg/m^3$ | double | 8 | Maximum chlorophyll concentration, as derived with the satellite |
| AOT_max (optional) | - | double | 8 | Maximum aerosol optical thickness at the longest NIR band, as derived with the satellite. |

#### 3.2.4.3 Output Data

In routine operation, the offline SVC Module provides a unique set of spectral gains (one per band) both in text file and in the dedicated ADF used by the processor in OMAPS (netCDF format). The online processor is then able to read this ADF and to apply the gains to the input Level1b radiometry. The impact of these gains can then be assessed outside the SVC Module on a validation dataset (Req-SVC-014); it is worth noting that gains notably impact the selection of the best AC algorithm in the RR-AC Module.

In test mode, the offline SVC Module also provides all individual gains, i.e., gains for all matchups, in the same format (one text file and one ADF per match-up). The module launches the online Level-2 processor on the individual scenes with the individual gains, and comparison of the marine reflectance with the in-situ data allows to check the relevance of these gains, on the calibration dataset.

*Table 3-27: Main properties of SVC output data set*

| | |
|---|---|
| ***Processing step*** | **SVC** |
| ***Status in processing*** | Output data |
| ***Naming convention*** | gain_vicarious |
| ***Storage*** | Text file + |
| ***Format*** | - Text file<br>- NetCDF for the ADF of the OC Processor containing gains (example for OLCI IPF: OL_2_ACP_AX.nc) |
| ***Temporal Coverage*** | Mission lifetime (for OLCI IPF: coverage described in name of the ADF) |
| ***Spatial Coverage*** | Global |
| ***Dimensions*** | Number of sensor bands (OLCI: 21 bands) |
| ***Size per file*** | - Text file: less than 1 KB<br>- ADF: related to size of the OC Processor (example for OLCI IPF: 2.7 GB) |
| ***Number of files*** | - Two in routine operation: text and netCDF format.<br>- Two per match-up (text+netCDF) in test mode |

*Table 3-28: Parameters in SVC output data set (text file)*

| Parameter | Unit | Data Type | Size (Bytes) | Description |
|---|---|---|---|---|
| gain_vicarious | dimensionless | float | 1 KB | List of sensor bands and associated mission averaged gains |

## 3.3 Validation and Quality Assessment

### 3.3.1 Overview

Validation and quality assessment of ocean colour products is a critical activity for the acceptance of products by users, whereby a key user the Copernicus Marine Service – Ocean Colour TAC is. Validation should be understood as comparison of the satellite derived products with a reference, while quality assessment also includes tests on stability of a quantity within a time series over an area which is known to be stable, or the comparison of a product derived from one sensor with a certain algorithm with the products derived from another sensor or a climatology.

The Validation (**VAL)** Module has the objective to support the scientific analysis of the output of the *ocean colour processor*. Thus, the requirements for this module concern validation methods and technical aspects concerning the output of the *ocean colour processor* and the interface to reference databases, in particular Copernicus Ocean Colour Reference database (OCDB).

The two basic methods, which were also applied in Ocean Colour CCI, are

- match-up comparison between reference data, measured in-situ, with the satellite derived ocean colour products
- Level 2 to Level 3 comparisons, i.e., the (daily) observation from given sensor with a temporally and/or spatially averages time series from another sensor or a climatology

The methods have been explained and results are documented in [RD-7]. This constitutes the basis for the validation and quality assessment requirements in the context of OMAPS. It includes match-up comparison of radiometric and water properties against reference measurements contained in the Ocean colour reference database as well as comparison of global trend analysis.

The VAL Module works on any bio-geo-physical variable which is provided in the ocean colour processor output, and which is also provided in the reference database. This comprises:

a) TOA radiances,

b) Rayleigh reflectances,

c) aerosol reflectances or similar, including the Rayleigh-aerosol contribution if relevant,

d) all transmittances,

e) reflectances related to surface and environment (e.g., sun glint, white caps),

f) aerosol (e.g., aerosol reflectances, Rayleigh-aerosol multi-scattering, aerosol model type),

g) quantities related to the decoupling between the water and the atmosphere or BPC,

h) BRDF values,

i) all pixel geometries.

The radiances and reflectances are provided at all sensor spectral bands except at those not suitable for Ocean Colour retrievals.

The VAL Module generates match-up statistics (numbers and plots) for the variable it is called. However, some of the statistics are already computed within the MATCHUP Module and stored in the MDB products. The statistics follow the specifications outlined in [RD-4]. It is possible to restrict the selection of matchups to spatial or temporal subsets, or other filter criteria. This allows analysis for certain sites, e.g., per Aeronet-OC site, analysis of temporal dependencies, and to filter e.g., on chlorophyll to select different water characteristics.

### 3.3.2 Input Data

#### *3.3.2.1 Diagnostic datasets*

The VAL Module includes a directory for auxiliary data where suitable satellite scenes for diagnostic sites are collected (“diagnostic dataset”). These satellite scenes cover diagnostic sites agreed with EUMETSAT, filled with Sentinel 3 OLCI A and B products.

#### *3.3.2.2 Output from OMAPS Processor modules*

As input for the VAL Module, the output from Online Processor and the various modules of the Offline Processor are used. These were already described in detail in previous sections:

- Ocean Colour L2 products from OMAPS Online Processor → section 3.1.9.1
- MDB products from MATCHUP Module → section 3.2.1.4
- MDB CSV summary files from MATCHUP Module → section 3.2.1.4
- Tables and plots from the RR-AC Module → section 3.2.2.4
- Tables and plots from the RR-IW Module → section 3.2.3.3

#### *3.3.2.3 Input to OMAPS Processor modules*

A common use case of the VAL Module is to invoke 'full' OMAPS workflows, e.g., to generate a set of MDBs with specific conditions and filter options for a certain validation activity, which is explicitly required in Req-VAL-010. In this case, a processing chain may need to start at the very beginning, i.e., with the generation of OC products from the Online Processor. In this sense, all inputs for the OMAPS modules required for the workflow also serve as input for the VAL Module. Again, all these inputs were already described in detail in the previous sections 3.1.2.1, 3.2.1.2, 3.2.2.3, and 3.2.3.2.

#### *3.3.2.4 Configuration for matchup scatter plots*

The VAL Module of the Offline processor provides a plot utility for the generation of sophisticated scatter plots of matching in situ / satellite variables (see section 3.3.3.1.2). A variety of properties of these plots can again be defined by the user via a configuration ini file. This configuration file is most conveniently generated by the Configuration GUI which is described in more detail in the SUM [AD-5]. Table 3-29 lists the available user options and their meaning, and **Error! Reference source not found.** shows an example for a configuration file used with a matchup test generation.

*Table 3-29: Configuration parameters used for validation scatterplots*

| PARAMETER | DESCRIPTION | POSSIBLE VALUES | DEFAULT |
|---|---|---|---|
| ***mdb_input_path*** | Full path to MDB NetCDF product | /path/to/file | |
| ***scatterplot_variables*** | Matchup variables to plot, comma separated | <xvar>, <yvar> | |
| ***plot_title*** | Plot title text | String | OMAPS Validation |
| ***plot_subtitle*** | Plot subtitle text | String | MDB Scatter Plot |
| ***plot_text*** | Text inside plot window | String | |
| ***plot_text_pos*** | Relative x and y position of text inside plot window, comma separated | any in [0.0,1.0], [0.0,1.0] | 0.7,0.3 |
| ***bg_color*** | Background RGB colour as hexadecimal code | | #7fffd4 |
| ***x_axis_label*** | Label on x axis | String | |
| ***y_axis_label*** | Label on y axis | String | |
| ***dataset_label*** | Label describing dataset | String | |
| ***x_min*** | Minimum of x data range to plot | Float | |
| ***x_max*** | Maximum of x data range to plot | Float | |
| ***y_min*** | Minimum of y data range to plot | Float | |
| ***y_max*** | Maximum of y data range to plot | Float | |
| ***log_scale*** | If set to True, logarithmic scale is used | True, False | False |
| ***legend*** | If set to True, a legend is drawn | True, False | True |
| ***legend_location*** | Relative x and y position of legend inside plot window, comma separated | any in [0.0,1.0], [0.0,1.0] | 0.05,0.5 |
| ***markers*** | Marker symbol from Python Matplotlib | | x |
| ***marker_size*** | Marker size in pt | Integer | 20 |
| ***markers_color*** | Marker RGB colour as hexadecimal code | | #0000ff |
| ***identity_line*** | If set to True, identity line is drawn | True, False | True |
| ***reg_line*** | If set to True, regression line is drawn | True, False | True |
| ***reg_line_color*** | Regression line RGB colour as hexadecimal code | | #ff0000 |
| ***confidence_limits*** | If set to True, confidence range, indicated as dashed lines, is drawn | True, False | False |
| ***confidence_limits_pct*** | Percentace value of confidence range | Integer in ]0,100[ | 95 |
| ***fig_text*** | Some additional descriptive figure text | String | |
| ***fig_text_pos*** | Relative x and y position of figure text inside plot window, comma separated | any in [0.0,1.0], [0.0,1.0] | 0.0,0.0 |
| ***save_plot_dir*** | Directory where plot is saved as png | /path/to/file | |
| ***save_plot_file*** | Plot file without extension (.png is added) | String | |

```
[MDB_SCATTERPLOT]
; configuration for validation
mdb_input_path = /home/olafd/omaps_test/mdb/MDB_S3A_OLCI_L2_OCDB_SACSO_TEST.nc
scatterplot_variables = insitu_rrs_443,satellite_Rrs_443
plot_title =
plot_subtitle = OLCI vs OCDB insitu: RRS 443nm
plot_text = some plot text...
plot_text_pos = 0.7,0.3
bg_color = #7fffd4
x_axis_label = OLCI SACSO RRS 443nm
y_axis_label = OCDB insitu RRS 443nm
dataset_label = OLCI vs. insitu
x_min = 0.0
x_max = 0.02
y_min = 0.0
y_max = 0.02
log_scale = False
legend = True
legend_location = 0.05,0.5
markers = x
marker_size = 20
markers_color = #0000ff
identity_line = True
reg_line = True
reg_line_color = #ff0000
confidence_limits = False
confidence_limits_pct = 95
fig_text = some figure text...
fig_text_pos = 0.19, 0.7
save_plot_dir = /home/olafd/omaps_test/val/sacso
save_plot_file = MDB_S3A_OLCI_L2_OCDB_RRS_sacso_RRS_443nm
```

*Figure 36: Example of a config ini file for a validation scatterplot.*

#### 3.3.2.5 Time series analysis

Exact matchups between satellite overpass and in-situ measurements are a small subset of the available in-situ data. However, if repeated measurements are taken at a given location, then timeseries analysis can be used to confirm whether the satellite products are correctly describing the phenology of the area, even if direct point to point comparisons are not possible. Well-known suitable sites for such analysis are the Hawaii Ocean Time-series (HOT) for chlorophyll-a, or MOBY (Marine Optical BuoY) for reflectances. Examples of such time series analysis are presented in the OMAPS PVER [AD-10].

The satellite input data for this time series analysis are the OLCI Level-1 TOA radiance products (section 3.2.1.2.1). From these, a list of files that overlap with the location of interest needs to be generated. This can e.g., be done with the OLCI granule finder, a utility tool which is provided with the OMAPS Offline Processor. It is described in more detail in the OMAPS Source Code Installation and Software User Manual [AD-6]. This subset of satellite products is processed by the Online Processor, applying the atmospheric correction of interest (e.g., POLYMER). This results in a set of OC L2 products from which the satellite for the comparison with the in situ data (taken from the Copernicus OCDB) are extracted. The data extraction as well as the time series plots (section 3.3.3.2) are realised with Python tools which are included in the VAL Module.

### 3.3.3 Output Data

#### 3.3.3.1 Matchup analysis and statistics

##### 3.3.3.1.1 Text output

Comprehensive text output of matchup analysis and statistics is already given by the MDB CSV summary files, generated by the MATCHUP Module, and described in more detail, together with an example, in section 3.2.2.4. This output would be generated by the MATCHUP Module itself, or in a validation workflow if invoked by the VAL Module as mentioned above.

##### 3.3.3.1.2 Scatter plots

As said above, the VAL Module of the Offline processor provides a plot utility for the generation of sophisticated scatter plots for any of the matching in situ / satellite variables. As input, a configuration file described in section 3.3.2.4 and a corresponding MDB NetCDF product containing the plot data are needed.

Figure 37 shows an example of a scatter plot which was generated as described above. The plot shows OLCI L2 (IPF) vs. OCDB in situ for Rrs(443nm). The region is around Hawaii (MOBY). The time interval is 20170201-20180430, but only a subset of 10 OLCI L2 products was used here as input. Plot features include a description text field, a regression line, identity line and various statistical parameters.

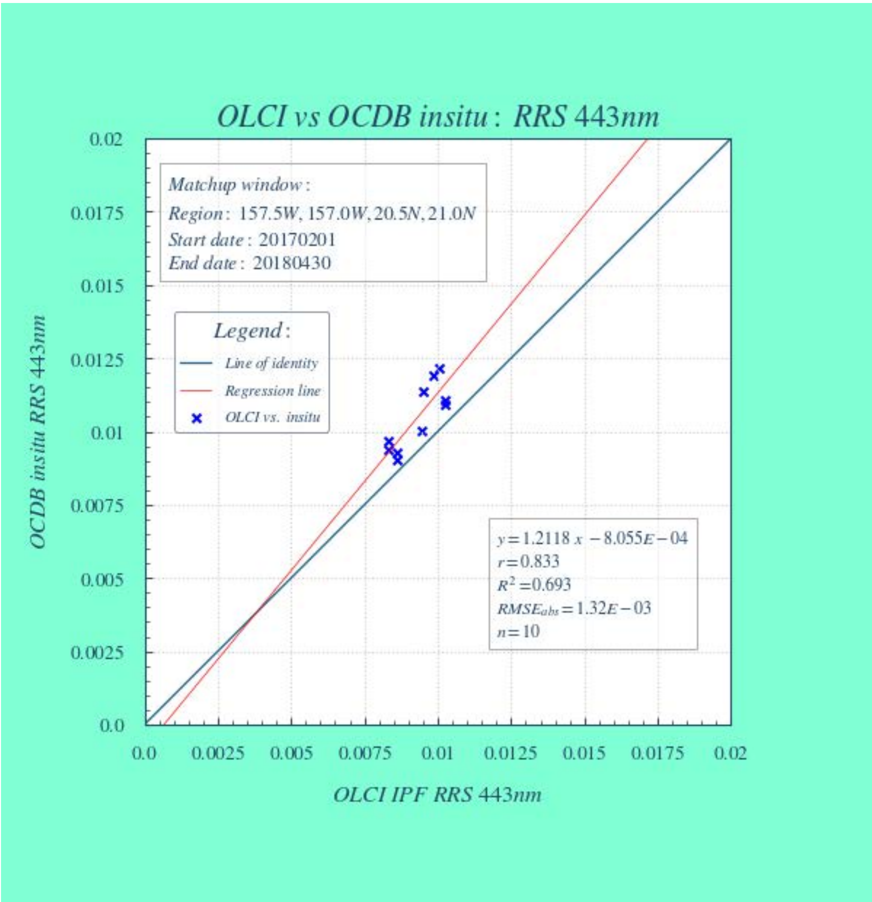



*Figure 37: Scatterplot of OLCI L2 (IPF) vs. OCDB in situ for Rrs(443nm). The region is around Hawaii (MOBY). Time interval is 20170201-20180430, but only a subset of 10 OLCI L2 products was used here as input.*

#### *3.3.3.2 Time series plots*

The output data of the time series analysis consists of plots such as the example shown in Figure *38*. Details and many more examples are given in the PVER [AD-10].

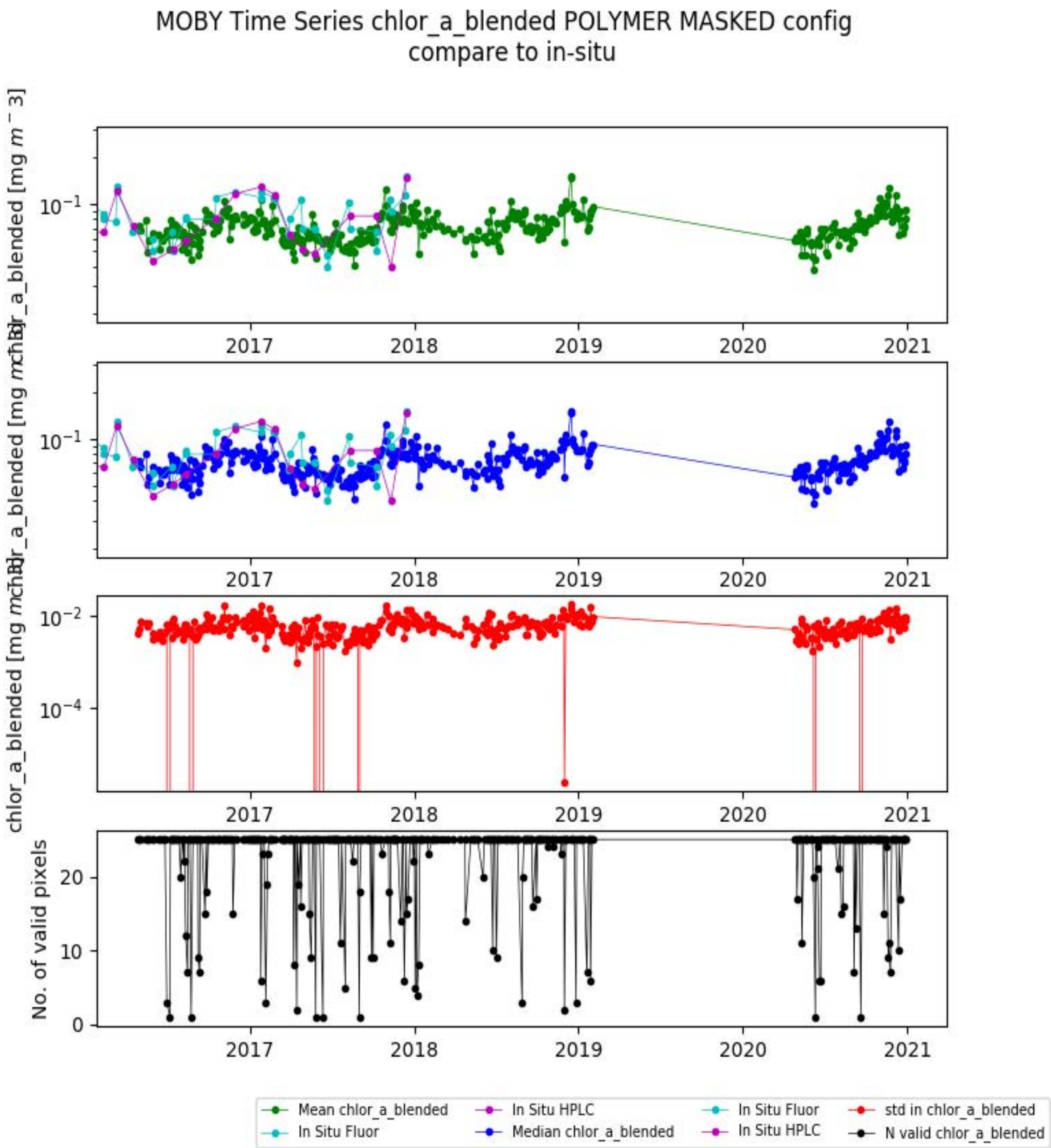



*Figure* 38*: Comparison of the extracted mean and median chlorophyll-a estimates for a 5x5 pixel area centered on the MOBY location, compared against HOTs in-situ measurements of chl-a. Taken from [AD-10].*