Ocean Colour Multi-Mission Algorithm Prototype System (OMAPS)

Algorithm Theoretical Baseline Document for Matchup Generation

Ref: D2.1.5
Date: 03.12.2021
Issue: 1.0
For: EUMETSAT
Ref: ID 995318

Table of Contents

Document Control 3
List of Acronyms 3
1 Purpose and scope 5
2 Introduction 5
3 Match-up generation input and output 7
3.1 Satellite input data 7
3.1.1 OLCI Level-1 TOA radiance products 7
3.1.2 OLCI Level-2 water products 7
3.2 In situ input data 7
3.2.1 Copernicus OCDB 7
3.3 Matchup configuration 7
3.4 Matchup variables 7
3.5 Intermediate products 7
3.6 Final products 8
4 S3 Matchup Protocol 8
4.1 Time difference between in situ and satellite measurements 8
4.2 Satellite data 8
4.2.1 Spatial window for extraction 8
4.2.2 BRDF correction 8
4.2.3 OWT determination 9
4.2.4 Per-pixel filtering criteria 9
4.2.5 Outlier removal 9
4.2.6 Validity of a macro pixel 10
4.3 In situ data 10
4.3.1 Band shifting of water reflectance 10
4.4 Matchup statistics 10
5 References 11

Document Control

Version No.	Release Date	Author/Contributor	Reason for issue
0.8-draft	2021-11-19	Olaf Danne	internal review
1.0	2021-12-03	Olaf Danne	Update after EUMETSAT review

List of Acronyms

Acronym	***Description***
ATBD	Algorithm Theoretical Baseline Document
BRDF	Bidirectional Reflectance Distribution Function
EUM	Eumetsat
ESA	European Space Agency
GUI	Graphical User Interface
IODD	Input/Output Data Description
IPF	Instrument Processing Facility
OC	Ocean Colour
OCDB	Copernicus Ocean Colour Reference database
OLCI	Ocean and Land Colour Instrument
OMAPS	Ocean Colour Multi-Mission Algorithm Prototype System
OWT	Optical Water Type
SACSO	Spectral matching Atmospheric Correction for Sentinel Ocean colour measurements
S3	Sentinel-3
SVC	System Vicarious Calibration (general) or System Vicarious Calibration module of the *offline processor* subsystem
TOA	Top of Atmosphere

INDEX OF FIGURES

Figure 1: High level workflow of the OMAPS matchup module (taken from [7]). 6

INDEX OF TABLES

Table 2-1: Main differences of OMAPS MATCHUP and EUMETSAT MDB standard workflows 6

1 Purpose and scope

The primary mechanism for assessing the quality of ocean colour products retrieved from satellite remote sensing data is through comparison with reference measurements of known quality. Such measurements are often called in-situ measurements or ground-truth measurements, although airborne or drone-borne measurements may also be used for this purpose. Naturally, reference measurements are not free from uncertainty sources, so should also be reported with their corresponding uncertainty estimates.

One goal of the OMAPS project is to provide a functionality to find matchups between the satellite and reference data. This functionality is realised by a specific software module, following the recommendations of the S3 matchup protocols [1]. The purpose of this document is to describe the subsequent steps implemented in this matchup module, in particular describing or at least pointing to the scientific algorithms behind it. These algorithms are, of course, of general nature, and can in principle be implemented in any ocean colour science context. However, the descriptions in this document will closely be related to the OMAPS context and its specific requirements. Thus, this document should not be regarded as a 'pure' ATBD (starting from a particular physical problem), as it also illustrates the technical aspects of the specific implementation following the needs of OMAPS. In the end, the document shall allow an understanding of the OMAPS matchup module on a sufficient level of detail required by an end user.

2 Introduction

The basic concept of match-up generation is the following: two data sets exist for reference ground data and 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 (as said above), 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 publications are [2] and [3] which contain a detailed description of the in-situ match-up process. An update has been recently published [4]. However, most relevant for OMAPS is [1], as it has been developed specifically for OLCI on Sentinel 3 under the coordination of EUMETSAT.

The match-up module for OMAPS contains the following main steps to generate a matchup:

1. Select a valid reference in situ measurement; this has been measured at time t_{ref} 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 OMAPS requirements. Defaults for the relevant parameters are set as defined in [1]. The architecture of the OMAPS matchup module is similar to EUMETSAT's MDB generation scheme as described in [5] 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 (probably the most common use case), but it is also possible to use L1 TOA radiance products to generate L1 matchups as e.g. needed by the OMAPS SVC module [6].

The high-level workflow of the OMAPS matchup module is illustrated in Figure 1.

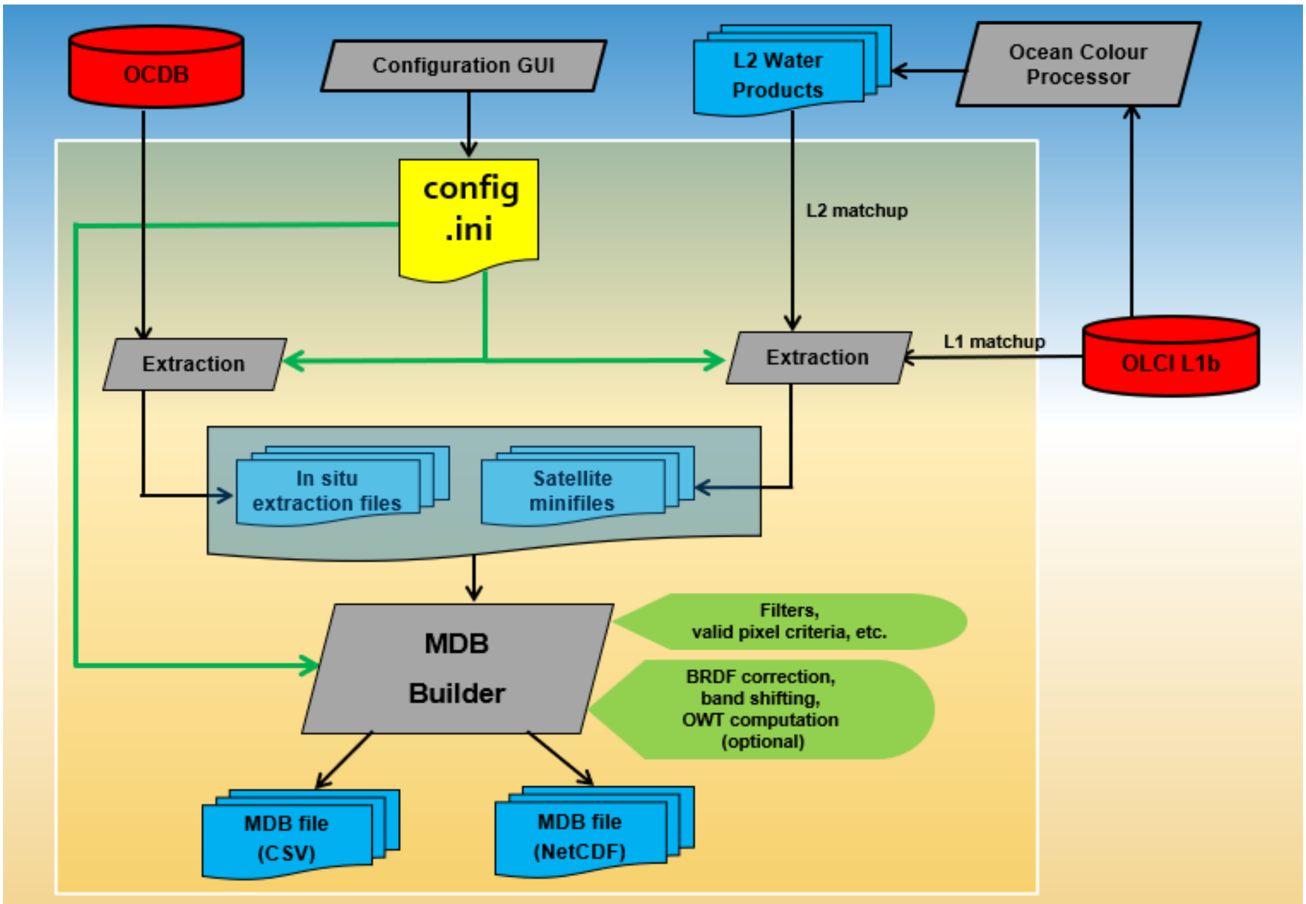

Figure 1: High level workflow of the OMAPS matchup module (taken from [7]).

Table 2-1 summarizes the main differences of the OMAPS matchup module workflow compared to the EUMETSAT MDB standard workflow (as described in [5], subsection 'Matchup Database Files')

Table 2-1: Main differences of OMAPS MATCHUP and EUMETSAT MDB standard workflows

	OMAPS MATCHUP MODULE	EUMETSAT MDB STANDARD
In-situ data source(s)	Copernicus OCDB	Copernicus OCDB, AERONET-OC, MOBY, SeaBASS
L2 Processors	OLCI Standard L2 (IPF), POLYMER, SACSO, L2GEN	OLCI Standard L2 (IPF)
Spatial window	1x1, 3x3, or 5x5 (configurable)	25x25
Temporal window for satellite extraction	configurable (in hours), including a 'time series' mode (infinite time window for satellite extraction)	+/- 24 hours
Variables considered for matchup and written to MDB	configurable	fixed sets, depending on in-situ data source, see [5]
Filter criteria	Highly configurable (see section 4.2 and [7])	default criteria from OLCI validation guidelines
MDB output formats	NetCDF, CSV summary file, or both (configurable)	NetCDF

3 Match-up generation input and output

For the generation of matchups, input data sources for both satellite and in situ data need to be available. In practice, the matchup generation needs to configurable to a large extent. For the OMAPS matchup module, this functionality is provided by a configuration file. The extraction of the relevant satellite and in situ data subsets results in various intermediate data. Finally, the matchup-products are provided in NetCDF and/or CSV format.

Input, intermediate, and final products of the OMAPS match-up generation are described in brief below. A much more comprehensive description with examples is given in the OMAPS IODD [7].

3.1 Satellite input data

In compliance with the requirements, the satellite data sets to be used for the matchup generation in OMAPS are 1) OLCI Level-1 TOA radiance products, and 2) OLCI Level-2 water products generated by the OMAPS Ocean Colour processor.

3.1.1 OLCI Level-1 TOA radiance products

The format and content of OLCI Level-1 TOA radiance products is described in full detail in [8]. 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.

3.1.2 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 [7].

3.2 In situ input data

3.2.1 Copernicus OCDB

For OMAPS the Copernicus OCDB is the required source for the in-situ measurements. 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 content and available variables can be different for the given datasets.

Among the variables available from the OCDB are remote sensing reflectance, concentration of chlorophyll-a, inherent optical properties such as phytoplankton absorption (a_{ph}), total absorption of detritus and gelbstoff (a_{dg}), backscattering of particles (b_{bp}) and the diffuse attenuation coefficient (K_d). Examples of datasets in the OCDB are shown in [7].

3.3 Matchup configuration

In the OMAPS matchup module, the configuration interface is provided as configuration file (ini format). For convenience, this configuration file can be generated by the OMAPS Configuration Tool, which is described in more detail in [9]. The content of the matchup configuration file, together with a full example, is described in detail in [7].

3.4 Matchup variables

In the matchup configuration, the user needs to define in particular the matchup variables to be extracted for the in-situ vs. satellite comparison. The extracted data for these variables are written in the final output products of the matchup generation. The rules how to define these variables in the configuration file for the OMAPS matchup module are described in detail in [7]and [9].

3.5 Intermediate products

Intermediate products generated from these input data sources are:

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

Examples of these intermediate products, which might be of interest and use in specific cases, are shown in [7].

3.6 Final products

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. Again, the content of these files, together with full examples, are described in detail in [7].

4 S3 Matchup Protocol

The implementation of the OMAPS matchup module basically follows the S3 matchup protocol as described in [1]. In this section, the main steps of the implementation are illustrated. The corresponding scientific algorithms applied in these steps are documented. For well-known and widely used algorithms, common references are just given instead. Emphasis is put on the steps of the implementation which deviate from or go beyond the S3 matchup protocol to fulfil specific OMAPS needs.

4.1 Time difference between in situ and satellite measurements

An in-situ measurement considered as valid reference measurement must have been measured within a defined time window around the time of the satellite observation. In the OMAPS matchup module, the length of this time window is configurable, the default length is +/- 1 hour. Moreover, a special mode of the match-up module allows the extraction of time series. If this mode is activated, the time information of the reference data is ignored (i.e. the length of the aforementioned time window is set to infinity). Given this, the in-situ reference measurement is paired with all available spatially coincident observations of the remote sensing data.

4.2 Satellite data

4.2.1 Spatial window for extraction

In the OMAPS matchup module, the spatial window ('macro pixel') for the extraction of satellite data around a reference measurement is configurable. The user can select between a window of 3x3 or 5x5 pixels, which will cover most use cases, but it is also possible to 'degenerate' the macro pixel to a standard pixel by selecting a window size of 1x1. The default size is 3x3 pixels.

4.2.2 BRDF correction

In the case that match-ups for marine reflectance are extracted, a BRDF normalisation factors should be applied to the reflectance obtained from both in situ and satellite, in case these were not previously applied to the data. In the case of OMAPS, the matchup module provides this functionality as an option, thus the user can decide, depending on the type of input products, if a BRDF correction shall be applied. Apart from this user option, the BRDF correction is only applied by default in case of OLCI standard L2 products considered as input ('IPF' processor mode, see [7]). For the other L2 processors supported by OMAPS (POLYMER, SACSO, L2GEN), the BRDF correction is already done at L2 processing stage.

The BRDF correction algorithm implemented in OMAPS is not fully outlined here, as it is described in detail in [10], [11] and [12]. In brief, mandatory inputs for the correction are horizontal wind components, satellite geometry (VZA, VAA), solar angles (SZA, SAA), chlorophyll concentration, aerosol optical thickness at 865nm. These quantities are taken from the satellite input products, and from static auxiliary files which are provided with the OMAPS matchup module package. Output of the algorithm are the BRDF correction factors per pixel for the following 11 OLCI wavelengths: 400, 412.5, 442.5, 490, 510, 560, 620, 665, 673.75, 681.25, and 708.75 nm.

4.2.3 OWT determination

In the case that match-ups for marine reflectance are extracted, a further specific requirement for the OMAPS matchup module is that the optical water type shall be determined and provided as output.

The division of waters into optical types is an established concept in marine sciences. *Morel and Pri*eur (1977) [13] distinguished two water types, those where bulk optical properties are dominated by phytoplankton (Case-1) and those where bulk optical properties are uncoupled from phytoplankton (Case-2). This early division led to the development of 'case-1' and 'case-2' algorithms which were tailored to produce the best results under a given set of assumptions about the optical nature of the waters in question. With the introduction of multidimensional clustering techniques applied to remote sensing data (see [14], [15]), this classification of waters can be refined. More than 10 optical water types (OWT) were identified in both open-ocean and inland-water environments in [16] and [17], respectively.

In the OMAPS matchup module, the functionality is implemented using the algorithm provided by *Moore et al* (2014) [15]. The resulting number of OWTs depends on the type of the L2 input water products. For example, if the water product was generated using the POLYMER algorithm, the OWT algorithm provides 14 water types. In any case, the optical water type determination in the matchup module is only used if OWT variables are not already included in the L2 water input products.

4.2.4 Per-pixel filtering criteria

NB: The default values detailed in this section correspond to [1].

4.2.4.1 Sun zenith angle

Within the satellite macro pixel window, a pixel will be filtered (observation regarded as invalid) if the sun zenith angle for this pixel is greater than a threshold specified by the user. The default threshold is 70 degrees.

4.2.4.2 View zenith angle

Within the satellite macro pixel window, a pixel will be filtered (observation regarded as invalid) if the view zenith angle for this pixel is greater than a threshold specified by the user. The default threshold is 60 degrees.

4.2.4.3 Validity of a satellite observation (obtained from a set of pre-defined quality flags)

Within the satellite macro pixel window, a single pixel is regarded as invalid the pixel is filtered because it is flagged according to one or more quality flags belonging to a pre-defined set. The default flag set for OLCI standard L2 products is detailed in [1]. These valid expressions can be specified by the user in the configuration file. For example, if OLCI standard L2 products are considered ('IPF' processor mode), and all water pixels which are not flagged as invalid or cloudy, the entry in the configuration file should look like

```
valid_pixel_expr = WQSF.WATER and not WQSF.INVALID and not WQSF.CLOUD
```

The default expressions are different for the different input products supported, as described explicitly and in detail in [7]. Also, a description and an example for a full configuration file are given there.

4.2.5 Outlier removal

In a subsequent step, further single pixel outliers in the macro pixel are removed (observation regarded as invalid) if the following condition applies for the pixel value x of the 'reference variable':

$$x > \mu + f_{\mathrm{var}} \cdot \sigma \quad \textbf{or} \quad x < \mu - f_{\mathrm{var}} \cdot \sigma$$

where μ is the mean and σ is the standard deviation of the values of the valid pixels (i.e. those having surpassed 4.2.4.1, 2 and 3) within the macro pixel. The value f_{var} is the 'tolerance factor' which can be configured by the user. The default value is 1.5. The 'reference variable' must be properly specified by the user, as outlined detail in [7] and [9].

4.2.6 Validity of a macro pixel

4.2.6.1 Number of valid pixels

A satellite macro pixel is regarded as valid if a certain fraction of pixels within the macro pixel window is still valid after all filter criteria detailed in 4.2.4. This threshold number can be configured by the user. The default value is 50%, which means at least 13 valid pixels in a 5x5 window, and 5 pixels in a 3x3 window.

4.2.6.2 Coefficient of variation

A satellite macro pixel is still regarded as valid if the 'Coefficient of variation' does not exceed a threshold which can be configured by the user. The default is 0.2 at the 560 nm band. The 'Coefficient of variation' is computed as:

$$\mathrm{CV} = \frac{\sigma}{\mu}$$

where again μ is the mean and σ is the standard deviation of the valid values of the 'reference variable' within the macro pixel. Here, valid refers to the remaining pixels after applying all the filter criteria mentioned above, and after removing the outliers.

4.3 In situ data

4.3.1 Band shifting of water reflectance

In case of marine reflectance, the in-situ instruments have different spectral characteristics than the space borne sensors (central wavelength, SRF). These differences can be minimised according to a technique called "band shifting" if bands are sufficiently similar, which is often the case (by purpose).

Following the S3 matchup protocol, the band shifting functionality implemented in the OMAPS matchup module uses the so-called Quasi-Analytical Algorithm (QAA), which was originally developed by *Lee et al*. (2002) [18]. The OMAPS implementation follows an updated version of the original algorithm, as described in detail in *Lee et al*. (2009) [19].

4.4 Matchup statistics

The following statistical quantities, defined in [1], are provided for a set of N valid matchups of any of the user defined matchup variables x (e.g. water reflectance, TOA radiance, etc.):

- Median Absolute Difference:

$$MdAD = \underset{1 \le i \le N}{\mathrm{median}} \left| x_{\mathrm{i,insitu}} - x_{\mathrm{i,sat}} \right|$$

where $\underset{1 \le i \le N}{\mathrm{median}}\{\dots\}$ represents the median over the set of N valid matchups.

- Median Absolute Percentage Difference:

$$MdAPD = \underset{1 \le i \le N}{\text{median}} \left| \frac{x_{\text{i,insitu}} - x_{\text{i,sat}}}{x_{\text{i,insitu}}} \right|$$

where $\underset{1 \le i \le N}{\text{median}}\{\dots\}$ represents the median over the set of N valid matchups.

In radiometry validations, spectral shape statistical analyses can bring additional useful information, in particular when comparing Level-2 OLCI standard products to any other algorithm products. Thus the OMAPS matchup module further provides the quantities:

- Spectral Angle Mapper:

$$SAM = \frac{1}{N} \sum_{i=1}^{N} \left(\operatorname{acos} \left(\frac{\langle Rrs_{\text{i,insitu}}, Rrs_{\text{i,sat}} \rangle}{\|Rrs_{\text{i,insitu}}\| \; \|Rrs_{\text{i,sat}}\|} \right) \right)$$

where $\langle Rrs_{\text{i,insitu}}, Rrs_{\text{i,sat}} \rangle$ is the dot product of Rrs vectors as derived in situ and from satellite along different spectral bands for each matchup *i*, and $\|Rrs_{\text{i,insitu}}\|$ and $\|Rrs_{\text{i,sat}}\|$ are the Euclidean norms of the same vectors.

- Chi-Square:

$$\chi^2 = \frac{1}{N} \sum_{i=1}^{N} \left(\sum_{\lambda} \frac{(Y(\lambda)_{\text{i,insitu}} - Y(\lambda)_{\text{i,sat}})^2}{Y(\lambda)_{\text{i,insitu}}} \right) \text{ with } Y(\lambda)_{\text{i}} = \frac{Rrs(\lambda)_{\text{i}}}{Rrs(560)_{\text{i}}}$$

5 References

[1]: EUMETSAT (2021): Recommendations for Sentinel-3 OLCI Ocean Colour product validations in comparison with in situ measurements – Matchup Protocols. Issue v7, Doc.No. EUM/SEN3/DOC/19/1092968, 17 May 2021; https://www.eumetsat.int/media/44087

[2]: 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.

[3]: 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

[4]: 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

[5]: EUMETSAT OCDB User Manual. Copyright 2019, Copernicus Revision 763d8907; https://ocdb.readthedocs.io/en/latest/index.html

[6]: Ocean Colour Multi-Mission Algorithm Prototype System (OMAPS): Algorithm Theoretical Baseline Document - System Vicarious Calibration. Issue 1.2, 29 July 2021.

[7]: Ocean Colour Multi-Mission Algorithm Prototype System (OMAPS): Input Output Data Description. Issue 2.0, 15 October 2021.

[8]: Product Data Format Specification – OLCI Level 1 & Level 2 Instrument Products. ESA Ref: S3IPF.PDS.004, Issue 1.10, 28 May 2015.

[9]: Ocean Colour Multi-Mission Algorithm Prototype System (OMAPS): Source Code Installation and Software User Manual. Issue 1.6, 14 July 2021.

[10]: Wang, Y. F., Li, X. W., Ma, S. Q., Yang, H., Nashed, Z., and Guan, Y. N. (2005). BRDF model inversion of multiangular remote sensing: ill-posedness and interior point solution method. Proceedings of the 9th International Symposium on Physical Measurements and Signature in Remote Sensing (ISPMSRS), Vol. XXXVI (pp. 328–330).

[11]: Morel, A., D. Antoine and B. Gentili (2002): Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function. Applied Optics 41, 6289-6306.

[12]: A. Morel and B. Gentili (1996): Diffuse reflectance of oceanic waters. III. Implication of bidirectionality for the remote sensing problem. . Applied Optics 35, 4850-4862.

[13]: Morel, André, Prieur, Louis (1977): Analysis of variations in ocean color, Limnology and Oceanography, 4, doi:10.4319/lo.1977.22.4.0709.

[14]: Moore, T.S., Campbell, J.W., and M.D. Dowell (2009): A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product, Remote Sensing of Environment, Volume 113, Issue 11, pp 2424-2430; https://doi.org/10.1016/j.rse.2009.07.016.

[15]: Moore, T.S., Dowell, M.D., Bradt, S., and A.R. Verdu (2014): An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters. Remote Sens Environ. 143:97-111. doi:10.1016/j.rse.2013.11.021.

[16]: Jackson, T., Sathyendranath, S., and F. Mélin (2017): An improved optical classification scheme for the Ocean Colour Essential Climate Variable and its applications, Remote Sensing of Environment, Volume 203, 2017, Pages 152-161, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2017.03.036.

[17]: Spyrakos, E., O'Donnell, R., Hunter, P.D., Miller, C., Scott, M., Simis, S.G.H., Neil, C., Barbosa, C.C.F., Binding, C.E., Bradt, S., Bresciani, M., Dall'Olmo, G., Giardino, C., Gitelson, A.A., Kutser, T., Li, L., Matsushita, B., Martinez-Vicente, V., Matthews, M.W., Ogashawara, I., Ruiz-Verdú, A., Schalles, J.F., Tebbs, E., Zhang, Y., and A.N. Tyler (2018): Optical types of inland and coastal waters. Limnol. Oceanogr., 63: 846-870.

[18]: Lee, Z., Carder, K. L., and R.A. Arnone (2002): Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. Applied Optics 41, 5755-5772.

[19]: Lee, Z., Lubac, B., Werdell, P. J., and R.A. Arnone (2009): An Update of the Quasi-Analytical Algorithm (QAA v5). Tech. rep., International Ocean Colour Coordinating Group (IOCCG); http://www.ioccg.org/groups/software.html.