PML | Plymouth Marine Laboratory

EUMETSAT

# Ocean Colour Multi-Mission Algorithm Prototype System (OMAPS)

# Algorithm Theoretical Baseline Document for In-Water Algorithm selection, processing and blending

**Ref:** D2.1.4
**Date:** 28/07/2021
**Issue:** 1.3
**For:** EUMETSAT
**Ref:** ID 995318

# Table of Contents

## Document Control

| Version No. | Release Date | Author/Contributor | Reason for issue |
|---|---|---|---|
| 0.4 | 16/01/20 | T. Jackson | Initial draft for review by EUMETSAT |
| 0.5 | 15/04/20 | T. Jackson | Update for PM3 following feedback from EUMETSAT |
| 1.1 | 16/05/2020 | T. Jackson | Updates following feedback after PM3 |
| 1.2 | 10/06/2021 | T. Jackson | Updates for final release (typographic corrections) |
| 1.3 | 12/08/2021 | T.Jackson | Updates for comments from EUMETSAT including merging of Blending and in-water algorithm ATBDs. |

## 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 D1.2 Requirement Baseline for Ocean Colour Processor |

## Reference Documents

**Bezdek**, J. C., Li, W. Q., Attikiouzel, Y., & Windham, M. (1997). A geometric approach to cluster validity for normal mixtures. Soft Computing, 1 ,166{179. URL: http://dx.doi.org/10.1007/s005000050019. doi:10.1007/s005000050019.

**Bricaud**, A., Babin, M., Morel, A., Claustre, H., 1995. Variability in the chlorophyll specific absorption coefficients of natural phytoplankton: Analysis and parameterization. Journal of Geophysical Research 100, 13,321{13,332.

**Ciotti**, A. M., Lewis, M. R., Cullen, J. J., 2002. Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. Limnology and Oceanography 47 (2), 404{417.

**Cohen**, J., Cohen, P., 1983. Applied multiple regression/correlation analysis for the behavioural sciences. L. Erlbaum Associates.

**Brewin**, R. J. W., Sathyendranath, S., Muller, D., Brockmann, C., Deschamps, P.-Y., Devred, E., Doerffer, R., Fomferra, N., Franz, B., Grant, M., Groom, S., Horseman, A., Hu, C., Krasemann, H., Lee, Z., Maritorena, S., Melin, F., Peters, M., Platt, T., Regner, P., Smyth, T., Steinmetz, F., Swinton, J., Werdell, J.,

III, G. N. W., 2013. The Ocean Colour Climate Change Initiative: III. A round-robin comparison on in-water bio-optical algorithms. Remote Sensing of Environment.
**Doney**, S. C., Lima, I. D., Moore, J. K., Lindsay, K., Behrenfeld, M. J., Westberry, T. K., Mahowald, N., M., G. D., Takahashi, T., 2009. Skill metrics for confronting global upper ocean ecosystem-biogeochemistry models against field and remote sensing data. Journal of Marine Systems 76, 95-122.
**Efron**, B., 1979. Bootstrap methods: another look at the jackknife. Annals of Statistics 7, 1-26.
**Efron**, B., Tibshirani, R. J., 1993. An Introduction to the Bootstrap. Chapman and Hall, New York.
**Franz**, B. A., Werdell, P. J., 2010. A Generalized Framework for Modeling of Inherent Optical Properties in Ocean Remote Sensing Applications. In: Ocean Optics XX, Anchorage, Alaska, 27th Sept.-1st Oct. 2010.
**Friedrichs**, M. A. M., Carr, M.-E., Barber, R. T., Scardi, M., Antoine, D., Armstrong, R. A., Asanuma, I., Behrenfeld, M., Buitenhuis, E. T., Chai, F., Christian, J. R., Ciotti, A. M., Doney, S. C., Dowell, M., Dunne, J., Gentili, B., Gregg, W. W., Hoep_ner, N., Ishizaka, J., Kameda, T., Lima, I., Marra, J., M_elin, F., Moore, J. K., Morel, A., O'Malley, R. T. O., O'Reilly, J. E., Saba, V. S., Schmeltz, M., Smyth, T. J., Tjiputraw, J.,Waters, K.,Westberry, T. K., Winguth, A., 2009. Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean. Journal of Marine Systems 76 (1-2), 113.
Garver, S. A., Siegel, D. A., 1997. Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation: 1. Time series from the Sargasso Sea. Journal of Geophysical Research 102, 18,607-18,625.
**Gohin**, F., Druon J. N. & Lampert, L. (2002) A five channel chlorophyll concentration algorithm applied to SeaWiFS data processed by SeaDAS in coastal waters, International Journal of Remote Sensing,23:8, 1639-1661, DOI: 10.1080/01431160110071879
**Gohin**, F., Loyer, S., Lunven, M., Labry, C., Froidefond, J.M., Delmas, D., Huret, M., Herbland, A., 2005. Satellite-derived parameters for biological modelling in coastal waters: illus- tration over the eastern continental shelf of the Bay of Biscay. Remote Sens. Environ. 95, 29–46.
**Hu**, C., Lee, Z., Franz, B., 2012. Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. Journal of Geophysical Research 117, C01011.
**Hu**, Chuanmin & Feng, Lian & Lee, Zhongping & Franz, Bryan & Bailey, Sean & Werdell, Jeremy & Proctor, Christopher. (2019). Improving Satellite Global Chlorophyll a Data Products Through Algorithm Refinement and Data Recovery. Journal of Geophysical Research: Oceans. 124. 10.1029/2019JC014941.
**IOCCG** (2019). Uncertainties in Ocean Colour Remote Sensing. Mélin F. (ed.), IOCCG Report Series, No. 18, International Ocean Colour Coordinating Group, Dartmouth, Canada. http://dx.doi.org/10.25607/OBP-696
**IOCCG** (2014) QAA v6 formulation. http://www.ioccg.org/groups/Software_OCA/QAA_v6_2014209.pdf
**Jackson**, T., Sathyendranath, S., Mélin, F., 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.
**Lee**, Z., Carder, K. L., Arnone, R. A., 2002. Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. Applied Optics 41 (27), 5755{5772.
**Lee**, Z., Carder, K. L., Mobley, C. D., Steward, R. G., Patch, J. S., 1998. Hyperspectral remote sensing for shallow waters. 1. A semianalytical model. Applied Optics 37 (27), 6329{6338.
**Lee**, Z., Lubac, B., Werdell, P. J., Arnone, R., 2009. An Update of the Quasi-Analytical Algorithm (QAA v5). Tech. rep., International Ocean Colour Coordinating Group (IOCCG) Online: http://www.ioccg.org/groups/software.html (assessed 10/02/2012).
**Lee**, Z., Hu, C., Shang, S., Du, K., Lewis, M., Arnone, R., and Brewin, R. (2013), Penetration of UV-visible solar radiation in the global oceans: Insights from ocean color remote sensing, *J. Geophys. Res. Oceans*, 118, 4241– 4255, doi:10.1002/jgrc.20308.
**Maritorena**, S., Siegel, D., and Peterson, A., (2002) Optimization of a semianalytical ocean color model for global-scale applications, Appl. Opt. 41, 2705-2714

**Moore**, T. S., Campbell, J.W., Dowell, M.D., (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.
**Moore** TS, Dowell MD, Bradt S, Verdu AR. (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
**Morel**, André, Prieur, Louis, (1977), Analysis of variations in ocean color, Limnology and Oceanography, 4, doi:10.4319/lo.1977.22.4.0709.
**Morel**, A., Antoine, D., 2011. MERIS Algorithm Theoretical Basis Documents (ATBD 2.9) - Pigment Index Retrieval in CASE 1 Waters (PO-TN-MELGS- 0005), Issue 4, July 2011. Tech. rep., Laboratoire d'Oceanographie de Villefranche (LOV), MERIS ESL, ACRI-ST.
NASA, March 2010. Ocean Color Chlorophyll (OC) v6. URL http://oceancolor.gsfc.nasa.gov/REPROCESSING/R2009/ocv6/
**OC-CCI** (2016) Jackson, T., Brewin, R.J.W. and Sathyendranath, S. OC-CCI PVASR PART2 In-Water https://esa-oceancolour-cci.org/sites/esa-oceancolour-cci.org/alfresco.php?file=2c5867b3-9fa9-4631-8ec6-01b83d8c9da4&name=OC-CCI-PVASR-PART2%20In-Water_29-01-16.pdf
**O'Reilly**, J.E., Maritorena, S., Siegel, D.A., O'Brien, M.C., Toole, D., Mitchell, B.G., Kahru, M., Chavez, F.P., Strutton, P., Cota, G.F. and Hooker, S.B., 2000. Ocean color chlorophyll a algorithms for SeaWiFS, OC2, and OC4: Version 4. *SeaWiFS postlaunch calibration and validation analyses, Part*, *3*, pp.9-23.
**O'Reilly**, J.E., Werdell, P.J. (2019) Chlorophyll algorithms for ocean color sensors - OC4, OC5 & OC6, Remote Sensing of Environment, Volume 229, Pages 32-47, https://doi.org/10.1016/j.rse.2019.04.021.
**Pope**, R., Fry, E., 1997. Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements. Appl. Opt. 36, 8710–8722.
**Sathyendranath**, S. , Cota, G. , Stuart, V. , Maass, H., and Platt, T.(2001) Remote sensing of phytoplankton pigments: A comparison of empirical and theoretical approaches, International Journal of Remote Sensing, 22:2-3, 249 273, DOI:10.1080/014311601449925
**Smyth**, T. J., Moore, G. F., Hirata, T., Aiken, J., 2006. Semianalytical model for the derivation of ocean color inherent optical properties: description, implementation, and performance assessment. Applied Optics 45 (31), 8116-8131.
**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 Tyler, A.N. (2018), Optical types of inland and coastal waters. Limnol. Oceanogr., 63: 846-870.
**Valente**, A., Sathyendranath, S., Brotas, V., Groom, S., Grant, M., Taberner, M., Antoine, D., Arnone, R., Balch, W. M., Barker, K., Barlow, R., Bélanger, S., Berthon, J.-F., Beşiktepe, Ş., Brando, V., Canuti, E., Chavez, F., Claustre, H., Crout, R., Frouin, R., García-Soto, C., Gibb, S. W., Gould, R., Hooker, S., Kahru, M., Klein, H., Kratzer, S., Loisel, H., McKee, D., Mitchell, B. G., Moisan, T., Muller-Karger, F., O'Dowd, L., Ondrusek, M., Poulton, A. J., Repecaud, M., Smyth, T., Sosik, H. M., Twardowski, M., Voss, K., Werdell, J., Wernand, M., and Zibordi, G.: A compilation of global bio-optical in situ data for ocean-colour satellite applications, Earth Syst. Sci. Data, 8, 235–252, https://doi.org/10.5194/essd-8-235-2016, 2016.
Werdell, P. J., 2009. Global bio-optical algorithms for ocean color satellite applications. Eos Trans. AGU 90 (1), 4.
**Werdell**, P. J., Franz, B. A., Bailey, S. W., Feldman, G. C., Boss, E., Brando, V. E., Dowell, M., Hirata, T., Lavender, S. J., Lee, Z., Loisel, H., Maritorena, S., M_elin, F., Moore, T. S., Smyth, T. J., Antoine, D., Devred, E., d'Andon, O. H. F., Mangin, A., 2013. Generalized ocean color inversion model for retrieving marine inherent optical properties. Applied Optics 52, 2019-2037.
**Zhang**, T., and F. Fell (2007), An empirical algorithm for determining the diffuse attenuation coefficient $K_d$ in clear and turbid waters from spectral remote sensing reflectance, *Limnol. Oceanogr. Methods*, **5**, 457– 462.

## Acronyms and Definitions

| Acronym | Definition |
|---|---|
| AC | Atmospheric Correction |
| ATBD | Algorithm Theoretical Baseline Document |
| CCI | Climate Change Initiative |
| CDOM | Coloured Dissolved Organic Matter |
| ESA | European Space Agency |
| GIOP | Generalized Inherent Optical Property |
| GSM | Garver-Siegel-Maritorena |
| HOPE | Hyperspectral Optimization Process Exemplar |
| HPLC | High Performance Liquid Chromatography |
| IOPs | Inherent Optical Properties |
| MERIS | Medium Resolution Imaging Spectrometer |
| NASA | National Aeronautics and Space Administration |
| OC-CCI | Ocean Colour Climate Change Initiative |
| OC-DB | Copernicus Ocean Colour Database |
| OMAPS | Ocean Colour Multi-Mission Algorithm Prototype System |
| OWT | Optical Water Type |
| QAA | Quasi-Analytical Algorithm |
| RMSE | Root Mean Square Error |
| RR | Round Robin |
| $R_{rs}$ | Remote Sensing Reflectance |
| SeaDAS | SeaWiFS Data Analysis System |
| SeaWiFS | Sea-Viewing Wide Field-of-View Sensor |
| SoW | Statement of Work |

# 1 Introduction

## 1.1 Scope of the document

Following atmospheric correction (AC) of remotely sensed ocean colour data to generate surface reflectance products, in-water algorithms are implemented to derive additional products of interest to the ocean colour community (such as chlorophyll-a concentration). The development of new or improved in-water algorithms remains an active field of research, necessitating the continued comparison of algorithms in an objective and repeatable manner.

This ATBD describes the various in-water algorithms available for use within the OMAPS processor, how the optimal algorithms for use are chosen and the approach used for algorithm blending within the 'Online Processor' subsystem developed in the OMAPS project. This is done in order to provide optimal product performance.

The selection of optimal in-water algorithms involves a round-robin comparison of the performance of bio-optical algorithms for use with remotely-sensed reflectance ($R_{rs}$) ocean colour data. Here, we describe the required input data and the algorithmic and statistical approaches used to score and rank algorithms. This methodology follows the work undertaken in the ESA Ocean Colour-CCI (OC-CCI) project (Brewin et al. 2013, OC-CCI 2016).

The objective scoring system compares algorithms in a multi-metric round-robin. The performance of each model is compared to the best performing or best possible performance, on a metric by metric basis before the overall performance is assessed. The procedure also makes use of fuzzy classification, assessing the algorithm performance within waters characteristic of each member of a water class set. As the algorithm assessment relies on matchups with in-situ data, a bootstrapping method varies the composition of the matched data to try and remove the impact of sampling bias on the results.

If differing algorithms are found to perform better in differing optical water classes then the blending approach can be used to provide a seamless transition between algorithms. This is also described within this document. As with the algorithm Round Robin, the approach is based upon the work performed by the Ocean Colour Climate Change Initiative (OC-CCI) project.

## 1.2 Blending Context

The requirement for algorithm blending results from the desire to create the highest quality products across a range of in-water conditions with a single processing chain. Following an assessment of the suitability of candidate algorithms for a range of optical conditions, the approach presented here allows the application and blending of multiple algorithms.

This reflects the state-of-the art as most algorithms are designed/parameterised to work for a given region, range in chlorophyll-a concentrations or water type. The approach presented should improve the overall performance of the final derived products and allow a reduction in product uncertainties.

Within the context of ocean colour, a successful algorithm blending should combine algorithms seamlessly to produce the optimal final product without introducing boundary artefacts. Work to achieve this goal was undertaken within the OC-CCI project, building upon the work of (Moore et al., 2014). The implementation and demonstrated utility of that approach is covered by Jackson et al. (2017) and the OC-CCI OCAB_ATBD. Here we will summarise the approach in general terms and describe its place and interactions within the OMAPS system.

Figure 1 shows how these stages (Round robin assessment and algorithm blending) are linked to provide blended output products from the OMAPS processor. The diagram is general in its construction, so it is worth noting that within the OMAPS system, Stage 1 is performed by the offline processor subsystem and Stage 2 is performed by the online processor subsystem.

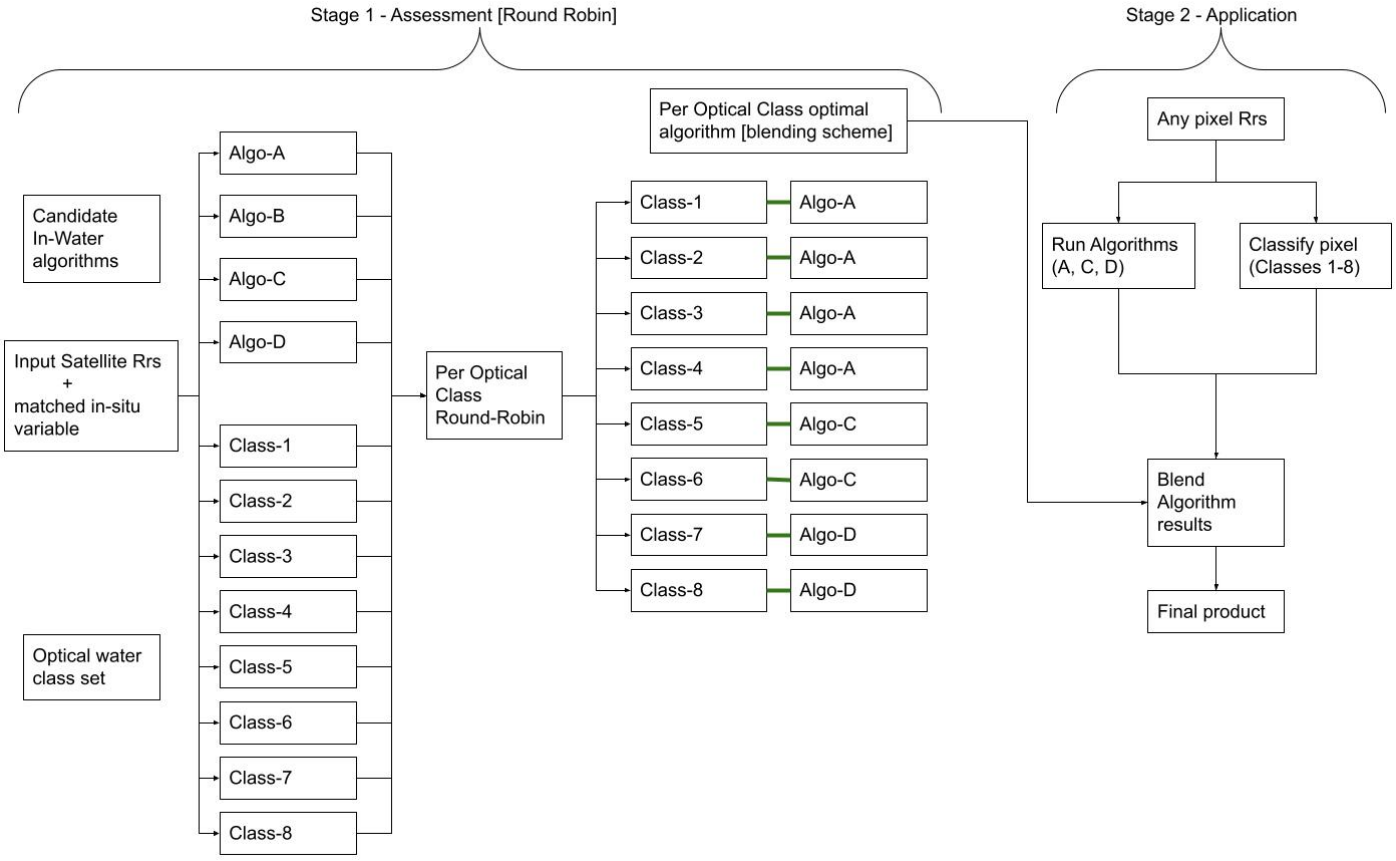


*Figure 1: Schematic diagram of the two stages involved in creating blended output products. Stage 1 is the assessment of algorithm performance across a range of optical types. Stage 2 is the application of the resulting blending scheme to Rrs data.*

# 2 Prerequisites

## 2.1 In situ database

A fundamental component of the algorithm analysis is the underlying in situ data base. In the case of OC-CCI, this is the titled "the Ocean Colour CCI database". It is a compilation and quality screening of existing open database (such as SeaBASS, MERMAID, Aeronet-OC) and proprietary measurements of the OC-CCI team and was published as Valente(2016). For OMAPS the Copernicus Ocean Colour Database (OC-DB) will be the source for reference measurements (https://ocdb.eumetsat.int). Within OC-CCI, the matchups were performed on the L3 data but within OMAPS we will use the L2 data processed using the OMAPS online processing system.

The following variables are of particular interest in the context of in-water algorithm validation: remote sensing reflectance, concentration of chlorophyll-a, inherent optical properties like phytoplankton absorption ($a_{ph}$), total absorption of detritus and gelbstoff ($a_{dg}$), backscattering of particles ($b_{bp}$) and the diffuse attenuation coefficient ($K_d$).

The in-situ data have to be homogenised, quality controlled and merged to commonly named variables. Metadata of each in situ measurement such as original source, cruise or experiment, and principal investigator has to propagated throughout the work and made available in the final data set. By making the metadata available, it is also possible to analyse each set of data separately and explore its impact on ocean colour validation.

## 2.2 Candidate in-water algorithms

The candidates for the in-water algorithm intercomparison have to be available (codified) for interaction with the RR-OC module. For the purpose of the RR-OC it is technically sufficient to have an executable of the code. However, for deeper understanding of the algorithm performance it is important to have open access to the source code. There is a large suite of potential products that can be generated from $R_{rs}$ data but in the context of OMAPS we will focus on algorithms designed to derive Chlorophyll-a or $K_d$ products. For a given round robin exercise, algorithms will be assessed on a product-by-product basis. This means that we will only assess performance in the context of chl-a, or $K_d(490)$ (or if you were looking at IOP matchups you would consider each IOP variable separately). The round robin module will assess products based on the product strings given in the config file. This means that you can compare any subset of empirical models or empirical + semi-analytical models provided that the products have been computed and extracted into the matchup products.

Generally, algorithms for ocean colour products are either semi-analytical or empirical. The following sections describe potential candidate semi-analytical models, usually designed to retrieve Inherent Optical Properties (IOPs), and the empirical chlorophyll algorithms. The listed set of algorithms were chosen to follow on from the OC-CCI activities. The empirical algorithms include both 2 band, 3 band and 4 band algorithms as though algorithms with more bands can make use of the additional spectral information (potentially helping accuracy) they also require a high level of accuracy and stability across a larger set of wavebands (O'Reilly and Werdell 2019). The mapping of optimal algorithm to optical waterclass is done on a statistical basis following the round robin multi-metric assessment, rather than a pre-defined assumption about which algorithms we might expect to perform best in each optical environment.

### 2.2.1 Chlorophyll-a models

Chlorophyll ($C$) algorithms incorporated into the comparison are described in the following section. Chlorophyll is a direct output from the models GSM and GIOP but not for semi-analytical models QAA, HOPE and PML. For the purposes of intercomparison, one can estimate chl from the QAA, and HOPE models as a function of $a_{ph}(443)$ using a power-law relationship (e.g. Bricaud et al. (1995)), such that

$$C = \left[\frac{a_{ph}(443)}{A}\right]^{\frac{1}{B}} \quad (1)$$

where, A and B are positive empirical parameters. Under OC-CCI the empirical parameters were computed using the in-situ database and set to A = 0.0549 and B = 0.7429.
It is worth noting that the empirical conversion from $a_{ph}(443)$ to chlorophyll is merely introduced to facilitate the comparison, it is not a feature of the original algorithms. In addition to the semi-analytical models, a variety of empirical chlorophyll algorithms were also incorporated into the comparison and are described below.

### 2.2.2 Chlorophyll-a from semi-analytical models through IOPs.

Semi-analytical models used in the comparison are described in this section. The term `semi-analytical models' will hereafter describe the following set of algorithms (2.2.2.1 to 2.2.2.4). However, we acknowledge that some of the models vary in their use of analytical and empirical solutions to solve for the IOPs. These semi-analytical models are used to compute the total absorption coefficient ($a$), combined absorption by detritus and coloured dissolved organic matter or gelbstoff ($a_{dg}$), absorption by phytoplankton ($a_{ph}$), total backscattering coefficient ($b_b$) and the ratio of phytoplankton absorption at 555 nm to that at 443 nm $a_{ph}(555)/a_{ph}(443)$ (see Table 1 for all notations used). The ratio $a_{ph}(555)/a_{ph}(443)$ was used in this comparison as an index of the spectral shape of the phytoplankton absorption coefficient (Sathyendranath et al., 2001; Ciotti et al., 2002). The ratio of 555 nm to 443 nm was chosen as these wavelengths typically represent the minimum and maximum of the phytoplankton absorption spectra. However, we acknowledge that ratios of other wavelengths could have also been used.

*Table 1:Variable Notation*

| Notation | Variable | Units |
|---|---|---|
| $a(\lambda)$ | Total absorption coefficient | $m^{-1}$ |
| $a_{ph}(\lambda)$ | Phytoplankton absorption coefficient | $m^{-1}$ |
| $a_{dg}(\lambda)$ | Dissolved and detrital absorption coefficient | $m^{-1}$ |
| $a_{tot}(\lambda)$ | Total absorption coefficient | $m^{-1}$ |
| $b_b(\lambda)$ | Total backscattering coefficient | $m^{-1}$ |
| $C$ | Chlorophyll concentration | mg $m^{-3}$ |
| $K_d(490)$ | Diffuse downwelling attenuation coefficient | $m^{-1}$ |
| $R_{rs}$ | Remote sensing reflectance | $sr^{-1}$ |
| $\theta$ | Solar zenith angle | Degrees |

#### 2.2.2.1 QAA model

The algebraic Quasi-Analytical Algorithm (QAA) of Lee et al. (2002) was designed to retrieve IOPs in optically-deep waters. The model inversion is based on two steps: the first involves partitioning water reflectance into $b_b$ and $a$ and the second decomposing $a$ into $a_{dg}$ and $a_{ph}$. The model is referred to as `Quasi-Analytical' as parts of the inversion are based on analytical, semi-analytical and empirical approximations. There have been a number of updates to the QAA algorithm over and here we refer to version 6 (as implemented in SeaDAS 7.5).

#### 2.2.2.2 HOPE model

In the physics-based Hyperspectral Optimization Process Exemplar (HOPE) model of Lee et al. (1998, 1999) $R_{rs}$ is modelled as a function of IOPs, and when influencing the $R_{rs}$ signal, bottom depth and bottom albedo. Unknowns are derived from non-linear optimisation. The spectral shape of bottom albedo is set to one of two options (one for sand, another for grass) which is selected based on the $R_{rs}$ spectrum. The phytoplankton absorption coefficients were constrained to lie within an upper and lower boundary (e.g. $0.002 < a_{ph}(443) < 1.0$ $m^{-1}$) which corresponds to a chlorophyll-a range of 0.0053 to 75.8 mg $m^{-3}$ (Lee et al, 1998, Brewin et al, 2013).

#### 2.2.2.3 GSM model

The semi-analytical Garver-Siegel-Maritorena (GSM) model was initially developed by Garver and Siegel (1997) and later updated by Maritorena et al. (2002). The GSM model retrieves simultaneous estimates of chlorophyll ($C$), $a_{dg}(443)$ and $b_{bp}(443)$ from $R_{rs}(\lambda)$, assuming an underlying bio-optical model and using non-linear optimisation. Global parameters of the bio-optical model were initially assigned based on simulated annealing on a global quasi-real dataset, which are then used in the non-linear optimisation routine. These include a fixed chlorophyll-specific phytoplankton absorption coefficient ($a_{ph}$), $S_{dg}$ and the slope of $b_{bp}$. The chlorophyll ($C$), $a_{dg}(443)$ and $b_{bp}(443)$ are first retrieved by fitting the bio-optical model to

the observed $R_{rs}(\lambda)$. IOPs at any wavelengths are then obtained using $C$, $a_{dg}(443)$ and $b_{bp}(443)$ and their specific shape function from the bio-optical model. The version of the model implemented here is parametrised and run according to SeaDAS.

#### 2.2.2.4 GIOP model

The Generalized Inherent Optical Property algorithm (GIOP) (Franz andWerdell, 2010;Werdell et al., 2013) model is designed as a test platform for algorithm development and was the result of a NASA IOP Algorithm Workshop (see Werdell (2009);Werdell et al. (2013)). Whereas the GIOP model offers the user freedom to specify different parameterisations and optimisation approaches, a preliminary configuration for GIOP is available which includes: an assigned $a_{ph}$ following Bricaud et al. (1995) but normalised by 0.055 $m^2$ $(mgC)^{-1}$; a spectral backscattering dependency following the QAA; a fixed spectral slope for $a_{dg}(\lambda)$ of 0.018 $nm^{-1}$; (Morel et al., 2002) f/Q ratio for zero Sun angle and zero view angle, where Q(λ) is the ratio of upwelling irradiance to upwelling radiance and f(λ) captures the net effects of variation in sea state, illumination conditions, and water column content; and Levenberg-Marquardt optimisation. The preliminary configuration is used as default. It is designed to retrieve spectral IOPs and chlorophyll, and it is worth noting that this preliminary configuration could be changed with time. All IOPs ($a_{dg}$, $a_{ph}$, $b_{bp}$, and $a_{dg} + a_{ph}$) were constrained to lie within -0.005 and 5 $m^{-1}$. Retrievals were excluded if the reconstructed $R_{rs}$ spectrum, between 412-555 nm, differed from the observed $R_{rs}$ spectrum by more than 33 %.

#### 2.2.2.5 NASA OC2

The NASA OC2 chlorophyll algorithm (O'Reilly et al., 2000) is a polynomial algorithm that relates the log-transformed ratio of remote sensing reflectances (X) to the chlorophyll concentration ($C$). The OC2 uses a two-band blue-green reflectance ratio where

$$X = log_{10}\left[\frac{R_{rs}(\lambda_{blue})}{R_{rs}(\lambda_{green})}\right] \quad (2)$$

and chlorophyll ($C$) is estimated according to Eq. 3. The coefficient a0 to a4 are sensor dependant, as are the wavelengths $\lambda_{blue}$ and $\lambda_{green}$ (closest wavelengths to 490 and 555 respectively). As an example, for SeaWiFS, the values are a0 = 0.2511, a1= -2.0853, a2 = 1.5035, a3 = -3.1747 and a4 = 0.3383 and $\lambda_{blue} =$ 490 and $\lambda_{green} = 555$ (NASA, 2010).

$$C = 10^{(a_0+\sum_{i=1}^{4} a_i X^i)} \quad (3)$$

#### 2.2.2.6 NASA OC3

The NASA OC3S chlorophyll algorithm (O'Reilly et al., 2000) is similar to the OC2 and OC4 algorithms. The OC3 is a polynomial algorithm that relates the log-transformed ratio of remote-sensing reflectances (X) to the chlorophyll concentration ($C$). The OC3S uses a three-band blue-green reflectance ratio where

$$X = log_{10}\left[\frac{\max\left(R_{rs}(\lambda_1), R_{rs}(\lambda_2)\right)}{R_{rs}(\lambda_{green})}\right] \quad (4)$$

The bands $\lambda_1$ and $\lambda_2$ are sensor dependant, being the bands closest to 443 and 490 respectively. Chlorophyll (C) is then estimated according to Eq. 3 where a0 to a4 are again sensor dependant and empirically derived.
NASA OC4
The NASA OC4 chlorophyll algorithm (O'Reilly et al., 2000) is a polynomial algorithm that relates the log-transformed ratio of remote sensing reflectances (X) to the chlorophyll concentration (C). The OC4v6 uses a four-band blue-green reflectance ratio such that:

$$X = log_{10}\left[\frac{\max\left(R_{rs}(\lambda_1), R_{rs}(\lambda_2), R_{rs}(\lambda_3)\right)}{R_{rs}(\lambda_{green})}\right] \quad (5)$$

where the bands $\lambda_1$, $\lambda_2$and $\lambda_3$ are also sensor dependant, being the bands closest to 443 and 490 and 510 respectively. Chlorophyll ($C$) is then estimated according to Eq. 3 where, a0 to a4 are empirically derived and are sensor dependant.

#### 2.2.2.7 OC4ME

The OC4ME is a MERIS chlorophyll band-ratio algorithm (Morel and Antoine, 2011). Like the OC4, it is a four-band polynomial algorithm that relates the log-transformed ratio of irradiance reflectance ($R$) to the chlorophyll concentration ($C$). It is important to note that unlike the OC4 algorithm, this algorithm uses the irradiance reflectance which is the subsurface reflectance ( $E_u(0-)/E_d(0-)$ ) rather than the above surface $R_{rs}$. When implemented in the OLCI baseline processing the irradiance reflectance and chlorophyll-a concentration are calculated in an iterative loop as the estimate of $R$ depends upon $C$ and the estimate of $C$ depends upon $R$. Though this iterative algorithm is used as the default in OCLI baseline processing, the algorithm code has not been provided to the OMAPS project. As such, it could be included in a round robin comparison (following matchup extraction) but cannot currently be easily added to the OMAPS processor as we do not have an executable binary. We have included a description here for completeness but for further information we refer readers to Morel and Antoine (2011).

#### 2.2.2.8 OCI

The OCI model refers to the chlorophyll algorithm of Hu et al. (2012) as modified for use by NASA. This empirical algorithm was designed to improve the estimate of chlorophyll ($C$) in the global ocean at concentrations <0.2 mg $m^{-3}$. For low chlorophyll concentrations (<0.15 mg $m^{-3}$), the algorithm uses a colour index (CI), which is defined as the difference between $R_{rs}$ in the green region of the visible spectrum and a reference formed linearly between $R_{rs}$ in the blue and red region of the visible spectrum. For high chlorophyll concentrations (>0.2 mg $m^{-3}$), the OCI conforms to the OC4 algorithm, and for concentrations >0.15 and <0.2 mg $m^{-3}$ a mixture of the colour index (CI) and the OC4 algorithm is used, allowing a smooth transition from the CI to the OC4 with increasing chlorophyll.

#### 2.2.2.9 OCI2

The OCI2 refers to an updated parameterisation of the OCI following the work of Hu et al (2019). This involves a new parameterisation of the CI coefficients and a blending between concentrations of 0.25 and 0.4 mgChl $m^{-3}$.

#### 2.2.2.10 OC5

The OC5 algorithm pays a particular attention to the effect of the suspended matters on the retrievals of the Chlorophyll-a from satellite radiances. It includes an empirical parameterisation relating absorption of CDOM and scattering of TSM to the 412 and 555 nm channels (Gohin et al., 2002, 2005).

### 2.2.3 $K_d$ models

Two $K_d$ models are provided within the OMAPS processor system. Both can be used to derive $K_d$(490) estimates from input level 2 $R_{rs}$ data.

#### 2.2.3.1 Kd_Lee

The first, following on from the OC-CCI heritage, is the $K_d$ model of Lee et al (2013):

$$K_d(\lambda) = (1 + m_0 * \theta_s) * a(\lambda) + \left(1 - \gamma * \eta_w(\lambda)\right) * m_1 * \left(1 - m_2 * e^{-m_3 * a(\lambda)}\right) * b_b(\lambda) \qquad (6)$$

As can be seen in equation 6, the $K_d$ algorithm requires estimates of absorption and scattering as input. Within the OMAPS processor these are calculated using the QAA v6 model (IOCCG, 2014). The QAA model is used as it was shown to perform best in a round robin intercomparison as part of the OC-CCI project (Brewin et al., 2015). The latest version of the QAA also takes account of raman scattering on the Rrs signal (Lee et al., 2013) where:

$$R_{rs} = \frac{R_{rs}^{T}}{1 + RF} \tag{7}$$

Where $R_{rs}$ is the remote-sensing reflectance without contributions from raman scatter, $R_{rs}^{T}$ is the remote sensing reflectance (including raman contribution) and $RF$ is a raman correction factor. RF is calculated following Lee et al. (2013) with the equation:

$$RF(\lambda) = \alpha(\lambda)\left(\frac{R_{rs}^{T}(440)}{R_{rs}^{T}(550)}\right) + \beta_1(\lambda)\left(R_{rs}^{T}(550)\right)^{\beta_2(\lambda)} \tag{8}$$

Where $\alpha$, $\beta_1$, *and* $\beta_2$ are were derived for a set of wavelengths using hydrolight simulations (see table 2 in Lee et al., (2013)) and are then interpolated to nearby wavelengths as required.

#### 2.2.3.2 Kd_Zhang

The second $K_d$ model codified in the $K_d$ module of the online processor system is the model of Zhang and Fell (2007). This second model computes $K_d(490)$ using a polynomial equation, switching coefficients based on a reflectance ratio:

$$K_d(490) = \begin{cases} 10^{-0.843-1.459X-0.101X^2-0.811X^3} \; if \; \frac{R_{rs}(490)}{R_{rs}(560)} \geq 0.85 \\ 10^{-0.094-1.302X-0.247X^2-0.021X^3} \; if \; \frac{R_{rs}(490)}{R_{rs}(560)} < 0.85 \end{cases} \tag{9}$$

Where $X = \log_{10}(\frac{R_{rs}(490)}{R_{rs}(560)})$ if $\frac{R_{rs}(490)}{R_{rs}(560)} \geq 0.85$ and $X = \log_{10}(\frac{R_{rs}(490)}{R_{rs}(560)})$ if $\frac{R_{rs}(490)}{R_{rs}(665)} < 0.85$. The original work of Zhang and Fell (2007) was parameterised using NOMAD data but used channel mapping, treating bands as interchangeable, between the wavelength sets of (555,560,565) and (665,670) so these equations are sometimes written elsewhere with 555 instead of 560 and 670 instead of 665, but no modification is required.

Following an OC-CCI round robin it was decided that a single $K_d(490)$ algorithm, that of Lee et al.(2013), was optimal. This means that no blending was required for this product and the default algorithm for $K_d$ in the OMAPS processor is set to the Lee algorithm.

# 3 OC Algorithm Round Robin procedure

## 3.1 Matchups and filtering

As with the associated atmospheric correction round robin, a points scoring classification is used for model comparison to rank objectively the performance of the candidate in-water algorithms. Each variable is tested independently. For a given variable, such as chl, multiple algorithms are used to process the $R_{rs}$ data to in-water product values following atmospheric correction of remotely-sensed ocean-colour data. These values are then matched from the level-2 processed data to the in situ database. The estimated variable is then compared with the corresponding in situ value using each statistical test and a score is assigned for each test ranging from zero to two.

These tests are described in the following sections. If an algorithm is not capable of estimating the variable, it was given zero points for that test.

The samples are only compared when the measured and estimated variables (if produced by the model) conformed to the following variable specific requirements, which represent extreme upper and lower boundaries fixed to avoid the influence of spurious results on the statistical properties:

- $C > 0.001$ and $< 200$ mg m$^{-3}$
- $K_d >$ aw (Pope and Fry, 1997) and $< 10.0$ m$^{-1}$

The lower boundary for $C$ is based on the absolute accuracy for HPLC detection if all protocols are strictly followed (Aiken et al., 2009). The exclusion of spurious results was conducted on a variable-by-variable basis.

## 3.2 Metrics and scoring

### 3.2.1 Statistical Tests

To test the performance of the bio-optical algorithms the following univariate statistical tests were adopted that are commonly used in comparisons between modelled and in situ data (Doney et al., 2009; Friedrichs et al., 2009). It is important to bear in mind that some variables (such as chl-a) are log-normally distributed at the global scale. This means that for those variables the statistics and scoring should be performed on log-transformed values to allow the statistical relationships to hold true.

The following variables can be assumed to be log-normally distributed at the global scale:

- Chlorophyll-a
- $a_{ph}$
- $K_d$
- $a_{dg}$
- $a_{tot}$

#### *3.2.1.1 Pearson correlation coefficient*

The correlation coefficient $r$ (also called Pearson's product moment correlation) is calculated according to:

$$r = \frac{1}{N-1}\sum_{i=1}^{N}\left[\frac{X_i^M - \left(\frac{1}{N}\sum_{j=1}^{N} X_j^M\right)}{\left\{\frac{1}{N-1}\sum_{k=1}^{N}\left[X_k^m - \left(\frac{1}{N}\sum_{l=1}^{N} X_l^M\right)\right]^2\right\}^{0.5}}\right]\left[\frac{X_i^E - \left(\frac{1}{N}\sum_{m=1}^{N} X_m^E\right)}{\left\{\frac{1}{N-1}\sum_{n=1}^{N}\left[X_n^E - \left(\frac{1}{N}\sum_{o=1}^{N} X_o^E\right)\right]^2\right\}^{0.5}}\right] \quad (6)$$

where, X is the variable and N is the number of samples. The superscript E denotes the estimated variable (from the algorithm) and the superscript M denotes the measured variable (from the in-situ dataset). Note that the Pearson correlation coefficient assumes a linear relationship between variables and normal distributions. The correlation coefficient may take any value between -1.0 and 1.0.

##### 3.2.1.1.1 Scoring

The $r$ test involved determining whether the $r$-value for each model was statistically lower than the model with an $r$-value closest to 1 (the best model). This was determined using the z-score. The z-score may be used to determine if two correlation coefficients are statistically different from one another (Cohen and Cohen, 1983).

Knowing the *r*-value for two respective models ($r_1$ and $r_2$, for model 1 and 2 respectively) and knowing the number of samples used to determine the r-values ($n_1$ and $n_2$) one can determine the z-score using the Fisher's r-to-z transformation.

Making use of the sample size employed to obtain each coefficient, these z-scores of each *r*-value ($z_1$ and $z_2$) can be used to compute the overall z-score (Cohen and Cohen, 1983), such that:

$$z_1 = 0.5 \log\left(\frac{1+r_1}{1-r_1}\right) \tag{7}$$

$$z_2 = 0.5 \log\left(\frac{1+r_2}{1-r_2}\right) \tag{8}$$

$$z_{score} = \frac{z_1 - z_2}{\left(\left[\frac{1}{n_1-3}\right]\left[\frac{1}{n_2-3}\right]\right)^{0.5}} \tag{9}$$

Having determined the z-score, this can be converted into a p-value assuming a normal distribution. For the model comparison, a two-tailed test is used and the score for each model is based upon the p-value as follows:

- 0 points if p-value for the model tested < 0.01 (i.e statistically different to the best model)
- 1 point if p-value for the model tested is ≥ 0.01 (i.e statistically similar to the best model)
- 2 points if p-value for the model tested is ≥ 0.05 (i.e statistically very similar to the best model). This is not cumulative with a point for being ≥ 0.01.

#### *3.2.1.2 Root Mean Square Error*

The absolute Root Mean Square Error (Ψ) is calculated according to:

$$\Psi = \left[\frac{1}{N}\sum_{i=1}^{N}(X_i^E - X_i^M)^2\right]^{0.5} \tag{10}$$

The scoring for RMSE is covered in section 3.2.1.4 and given that Ψ and Δ are strongly related and the multi-metric approach should not score both of these metrics, only one or the other, and the scoring approach is the same for both. Given that RMSE is effectively composed of the same information contained within the bias and the unbiased-RMSE then we will not use RMSE in the scoring scheme.

#### *3.2.1.3 Bias*

The bias (δ) between model and measurement can be expressed according to:

$$\delta = \frac{1}{N}\sum_{i=1}^{N}\left((X_i^M - X_i^E)\right) \tag{11}$$

##### 3.2.1.3.1 Scoring

The closer the model bias (δ) is to the reference value of zero implies that the model corresponds well with the in situ data. However, a model could have a δ close to the reference value of zero, when compared with another model, but have a much larger 95 % confidence interval, implying lower confidence in the retrieved δ. Therefore, the following points classification was introduced for the bias:

- 0 points = the 95 % confidence interval of δ for a particular model is higher than the 1.5 times the smallest 95 % confidence interval of any model. In addition to this, the modulus of the bias ± its 95 % confidence interval did not overlap with zero.
- 1 point = either, the 95 % confidence interval of δ for a particular model less than 1.5 times the smallest model 95 % confidence interval, or the modulus of the bias ± its 95 % confidence interval overlaps with zero, but not both cases.
- 2 points = the 95 % confidence interval of δ for a particular model is less than 1.5 times the smallest model 95 % confidence interval and the bias ± its 95 % confidence interval overlaps with zero.

For both bias and centre-pattern Root Mean Square Error statistics the 95% confidence interval is computed from the standard error of the mean and the t-distribution of the sample size such that:

$$Confidence\ Interval = t_{0.025,n-1}\frac{S_n}{\sqrt{n}}\,, \tag{12}$$

Where $S_n$ is the standard deviation of the error, n is the number of matchups and t is the two-tailed t-distribution.

#### *3.2.1.4 Centre-pattern Root Mean Square Error*

The absolute centre-pattern (or unbiased) Root Mean Square Error (Δ) is related to the bias and rmse as:

$$\Delta^2 = \Psi^2 - \delta^2 \tag{13}$$

It describes the error of the estimated values with respect to the measured ones, regardless of the average bias between the two distributions.

##### 3.2.1.4.1 Scoring

In addition to computing Ψ and Δ for each model, it is possible to determine the confidence levels in the Ψ and Δ, which provide an indication of how confident one is in the statistics. The confidence levels are computed from the standard error of the mean percentage and the t-distribution of the sample size. Confidence levels provide a powerful way of highlighting differences and similarities between models. If the confidence intervals of two or more models overlap, then it can be assumed that the models have a statistically similar Ψ or Δ at the given confidence interval. For each model, the 90 % and 95 % confidence intervals are computed for both Ψ and Δ. Points for each model are awarded according to:

- 0 points = Ψ or Δ for the model tested is statistically higher than the Ψ or Δ for the best model (no overlap with a p-val of 0.01).
- 1 point = Ψ or Δ for the model tested is statistically similar to the Ψ or Δ for the best model (overlap with a p-val of 0.01 ).
- 2 points = Ψ or Δ for the model tested is statistically very similar to the Ψ or Δ for the best model (overlap with a p-val of 0.1.

As mentioned in section 3.2.1.2 we should not use both Ψ and Δ in the scoring scheme as we do not want to overweight or double score a single metric of performance. We have chosen to use bias (δ) and unbiased RMSE (Δ) as the RMSE (Ψ) can be derived from the other two statistics and using both δ and Δ gives a separation between metrics relating to accuracy and precision.

#### *3.2.1.5 Slope (S) and Intercept (I) of a Type-2 regression*

The performance of a model with respect to in situ data can be tested using linear regression between the estimated variable (from the model) and the measured variable (in situ data), such that:

$$X^E = X^M S + I \tag{14}$$

A slope (S) close to one and an intercept (I) close to zero is an indication that the model compares well with the in situ data. Type-1 regression typically assumes the dependent variable (in situ data) is known infinitely well, when in reality the in situ data are also affected by uncertainties (e.g. problems with in situ data sampling techniques) that are difficult to quantify. Therefore, we use a Type-2 regression, which instead of minimising the vertical distance between independent data and linear fit (as in Type-1 regression), minimises the perpendicular distance between independent data and linear fit.

#### 3.2.1.5.1 Scoring

In addition to computing the intercept (I) and the slope (S) from Type-2 regression, it is possible to compute the standard deviation on I and S. The closer the intercept (I) is to the reference value of zero and the closer the slope (S) is to the reference value of one, the better the fit between variables. Similar to the Bias, a model could have an intercept closer to the reference value of zero and a slope closer to the reference value of one, when compared with another model, but have a much larger standard deviation on its retrieved parameters, implying lower confidence in the fit. Therefore, to account for both these possibilities the following points classification is introduced for the slope (S) parameter:

- 0 points = the standard deviation of the S parameter for a particular model is higher than 1.5 times the smallest standard deviation in S. In addition to this, the S parameter ± its standard deviation does not overlap with 1.
- 1 point = either, the standard deviation of the S parameter for a particular model is less than 1.5 times the smallest standard deviation or S ± its standard deviation overlaps with 1, but not both cases.
- 2 points = the standard deviation of the S parameter for a particular model is less than 1.5 times the smallest standard deviation, and, the S parameter ± its standard deviation overlaps with 1.\

The following points classification was introduced for intercept (I) parameter:

- 0 points = the standard deviation of the I parameter for a particular model is higher than 1.5 times the smallest standard deviation in I. In addition to this, the I parameter ± its standard deviation does not overlap with that zero.
- 1 point = either the standard deviation of the I parameter for a particular model is less than 1.5 times the smallest standard deviation, or the I parameter ± its standard deviation overlaps zero, but not both cases.
- 2 points = the standard deviation of the I parameter for a particular model is less than 1.5 time the smallest standard deviation and the I parameter ± its standard deviation overlaps with zero.

### *3.2.1.6 Percentage of possible retrievals*

Considering that algorithms chosen for climate studies should perform well routinely, and globally, and should not be a source of more gaps in the data than would be the case if other algorithms were used, the percentage of possible retrievals (η) is an important criterion that should be considered in the comparison, calculated according to:

$$\eta = \frac{N^E}{N^M} 100 \tag{15}$$

where $N^E$ represents the number of retrievals using the model and $N^M$ represents the number of available matched in situ data points.

#### 3.2.1.6.1 Scoring

To compare the percentage of possible retrievals (η) between models, highest percentage and standard deviation of retrievals for all models was used. The following points criteria were set-up:

- 0 points = η of a model lies further than 1 standard deviation below the maximum algorithm η.
- 1 point = η of a model lies within 1 standard deviation of the maximum η.
- 2 points = η of a model is equal to the maximum η.

## 3.3 Bootstrapping and multi-metric scoring

To rank the performance of each model with reference to a particular variable, all points were summed over the set of statistical tests used. The total score for each model was then normalised to the highest score for a single model. A score of one indicates the model scored the highest total number of points across all tests, with values less than 1 showing the points score of each model relative to the highest achieved. Figure 2 shows a flow-chart illustrating the methodology of the scoring system used to compare models.

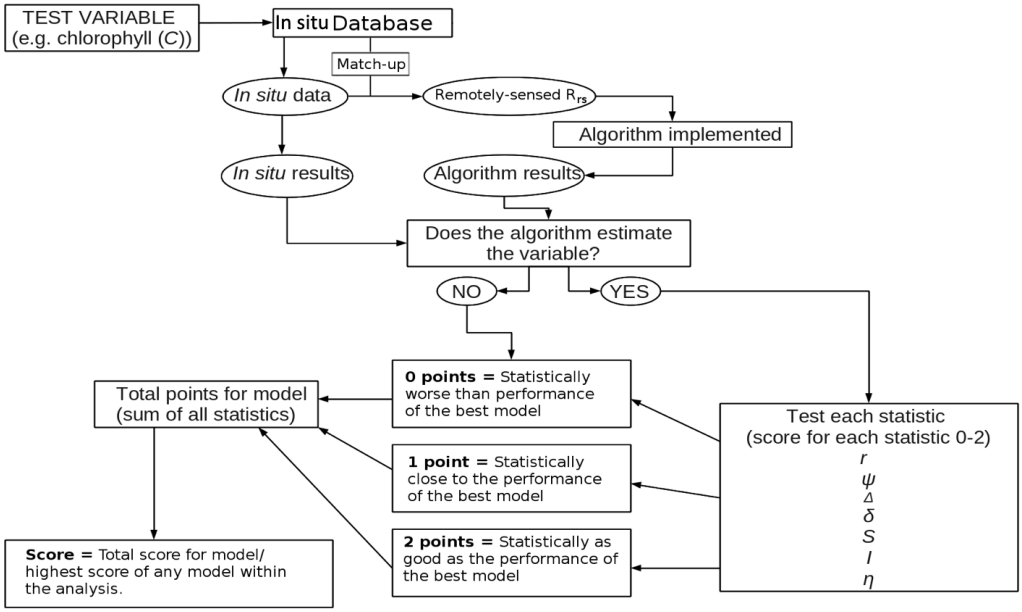


*Figure 2: Flow chart of multi-metric scoring approach for Ocean Colour algorithm comparison*

The stability of the scoring system, and the sensitivity of the scores, was tested using the method of bootstrapping (Efron, 1979; Efron and Tibshirani, 1993). This involves using sampling with replacement to randomly re-sample the in situ data to create 1000 new datasets, each the same size as the original dataset. The quantitative statistical methodology was then re-run for each new dataset (Monte-Carlo approach) and from the resulting distribution of scores, a mean score for each model was computed. Additionally, a 2.5 % and a 97.5 % interval on the bootstrap distribution is taken and assumed to be the confidence limits on the mean score for each model, rather than standard deviations on the bootstrap distribution, to avoid misinterpretation of results should the bootstrap distribution not follow a normal distribution or be skewed, for instance from the presence of outliers in the data.

## 3.4 Division by water class

The division of waters into optical types is an established concept in marine sciences. Morel and Prieur (1977) distinguished two water types, those where bulk optical properties are dominated by phytoplankton and (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 (Moore et al 2009, 2014), this classification of waters has become non-binary with >10 optical water types (OWT) identified in both open-ocean (Jackson et al., 2017) and inland-water (Spyrakos et al., 2018) environments.
The initial comparison and scoring of algorithms uses all the available in situ data. This in situ database has been collected in a variety of environments from open-ocean, clear waters to coastal environments with complex optical properties. In spite of a great deal of effort it is known that algorithms optimised to work in very clear waters are likely to perform worse in highly turbid waters, especially when compared to algorithms tailored for such environments. In order to assess the performance of the algorithms across a range of optical conditions the matched in situ and $R_{rs}$ data are divided based on the dominant optical water class (for example, those defined in the OC-CCI processing chain). Each optical class subset was then run through the same scoring processing as the total dataset.

# 4 Blending Algorithm description

## 4.1 Blending outline

The product blending processor will sit within the ocean colour processing subsystem. It must be used after the $R_{rs}$ spectra are assigned fuzzy memberships (Bezdek, 1997) to a pre-defined set of optical water classes (Jackson et al. 2017). It is also a prerequisite that an in-water algorithm round robin has been performed for the candidate product algorithms, such as those performed as part of the OC-CCI production process, in order to identify the optimal product algorithm per for each water class. This should be possible using the round-robin modules included in the OMAPS offline processor system.

Once waterclass memberships and per-waterclass optimal algorithms are known, one can apply equation 1 to calculate the weighted product estimate per pixel using $n$ waterclasses:

$$P_w = \frac{\sum_{i=1}^{i=n} P_i W_i}{\sum_{i=1}^{i=n} W_i}, \quad (1)$$

where $P_w$ is the weighted product value, $P_i$ is the product estimated using the optimal algorithm for water class $i$ and $W_i$ is the pixel membership to water class $i$. This is the same as the weighting scheme used by Moore et al (2014) and Jackson et al (2017).

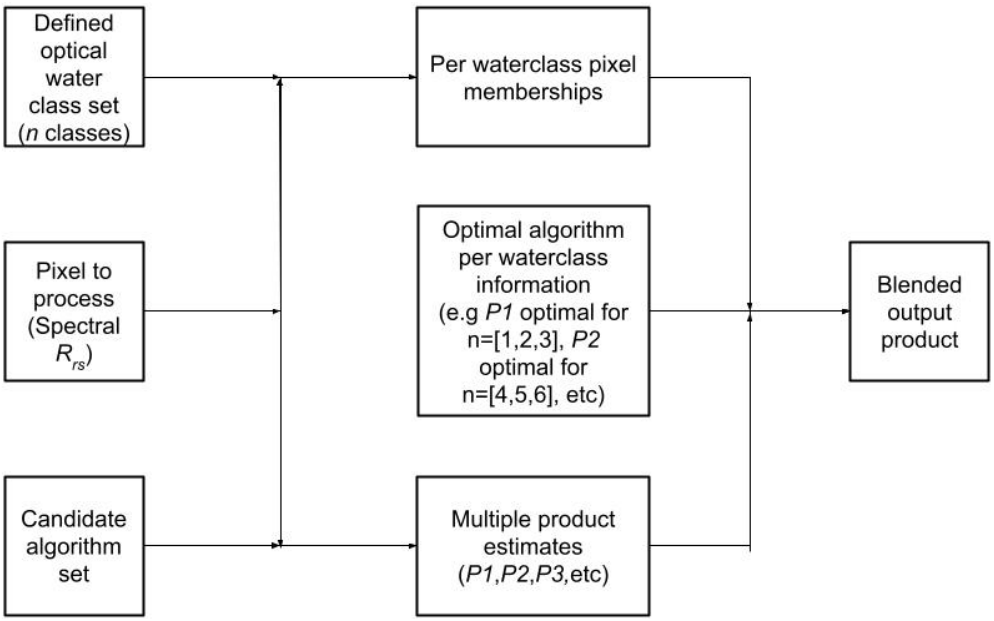


*Figure 3: Schematic diagram of application of algorithm blending to pixel spectrum*

## 4.2 Blending inputs

As shown in Figure 3, the blending algorithm requires a single source of data as input along with at least two predefined candidate algorithms, an optical waterclass set and information relating to the performance of candidate algorithms per water class. The input data is spectral $R_{rs}$ which should contain a bandset that is sufficient for both the comparison to the waterclass set and use with all candidate algorithms.

As an example, we might consider a blended chlorophyll-a product. P1, P2, P3 could be the OC4v6 NASA algorithm, a neural network derived Chl-a estimate and Chl-a from a semi empirical algorithm respectively.

The waterclass set used in the default implementation of the initial OMAPS processor for OLCI processing is shown below in Figure 4 and was created as part of the OC-CCI v5.0 processing for use with a MERIS/OLCI waveband set. This waterclass set can likely be improved upon for OLCI specific processing but was chosen as an initial set as it has been used operationally with an OLCI data stream. Future sets might be non-normalised to the spectral integral, make use of more of the OLCI wavebands and contain a different number of classes. Once a new waterclass set is created it can be easily called upon by editing the relevant configuration file.

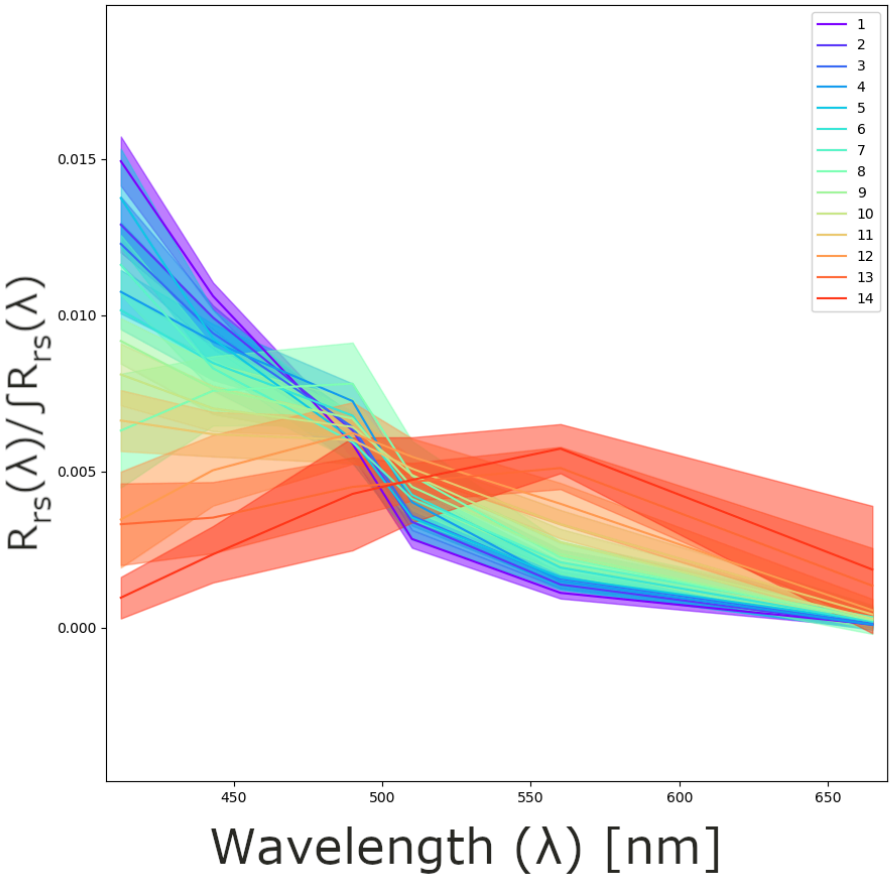


*Figure 4: Example Optical water class set using 6 bands and consisting of 14 clusters.*

## 4.3 Blending outputs

The result of running the blending processor is that a single additional variable is added to the netCDF file (or data structure). A single blended chlorophyll-a variable can be created from any given set of algorithms, set within a configuration file. In the OMAPS offline processor the default name for the output blended chlorophyll-a product is "chlor_a_blended".

## 4.4 Algorithm demonstration

As a demonstration of the application of this blending technique we have processed a granule (S3A_OL_1_EFR____20180716T073803_20180716T074103, covering the Red Sea), shown in Figure 5, using 8 candidate chlorophyll-a algorithms and blended four of them (OC2, OCx, OCI, and OCI2) into a blended chlorophyll-a product. Keep in mind that the OCI, OCI2 and OCx algorithms are themselves compound algorithms that switch between different empirical relationships. The top middle panel of Figure 5 shows the percentage increase or decrease in the blended product compared to the OCI2 only product. There is little difference in much of the clear waters of the Red Sea but as the waters become more chlorophyll-a rich in the north east and south east areas of the image we can see that the blended product is increased relative to the OCI2 estimates as the other algorithms begin to increase in weighting. The weighting factors for wach of the 4 contributing algorithms is shown in Figure 6.

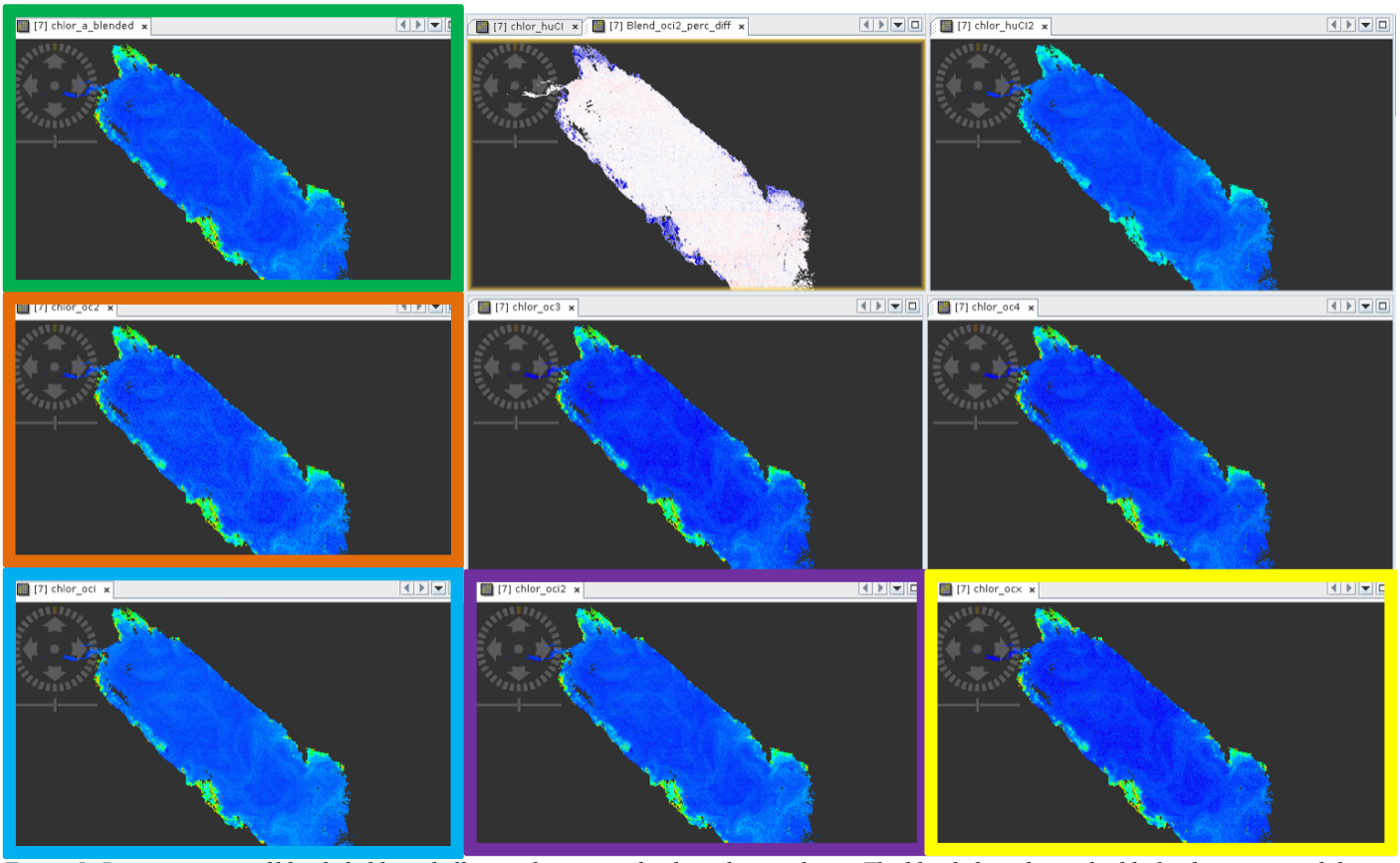


*Figure 5: Demonstration of blended chlorophyll-a product to single algorithm products. The blended product is highlighted in green and the contributing algorithms are the OC2, OCx, OCI and OCI2 (in orange, yellow, blue and purple respectively). The top middle panel shows the difference between the blended product and the HuCI2 algorithm (with blue being a positive difference, red negative and small differences in white (red and blue colour scales capped at ±40%).*

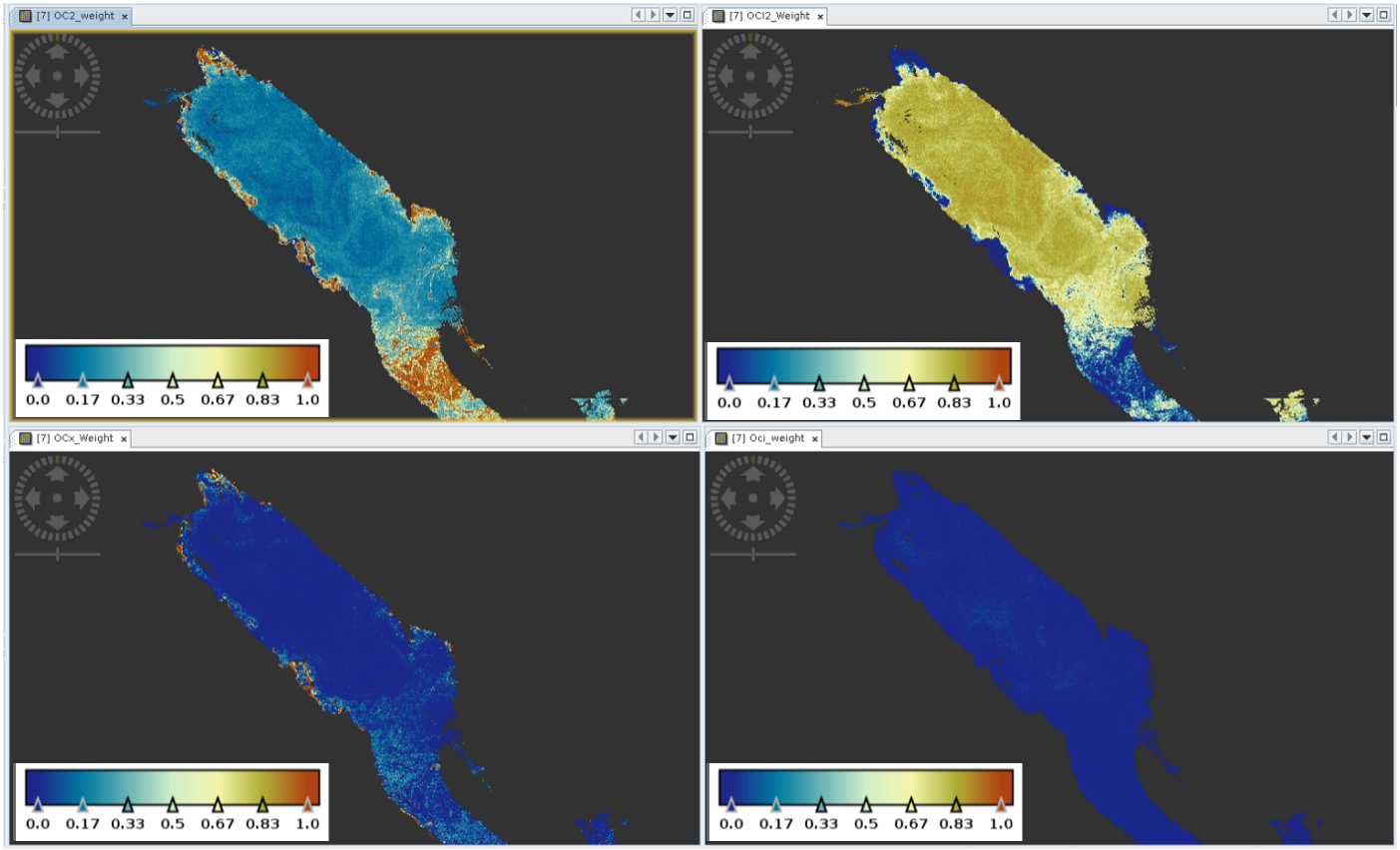


*Figure 6: Weighting factors for OC2, OCI2, OCx and OCI algorithms as used in the blending for Figure 5. Weighting factors are derived from summed subsets of waterclass memberships relative to total water summed waterclass memberships.*

## 4.5 Additional considerations

### 4.5.1 Product Uncertainties

In an ideal system there would be a comprehensive error propagation, beginning with known instrument performance, to provide final product uncertainties. This is not yet available for within ocean colour remote sensing (IOCCG 2019), though efforts are ongoing. An alternative approach used by the OC-CCI project is to treat the per-variable uncertainties of root mean square difference (RMSD) and bias as products in their own right, to be blended in the same manner as the chlorophyll-a estimates. In this instance one can use the waterclass membership weighting as given in equation 1, but the candidate algorithm estimates are replaced by 'per-waterclass' uncertainty statistics. An example for pixel bias (δ) is given in equation 2:

$$\delta = \frac{\sum_{i=1}^{i=n} \delta_i W_i}{\sum_{i=1}^{i=n} W_i}, \quad (1)$$

where $\delta_i$ is the waterclass specific bias estimate, derived from matched in-situ data. This is possible using the same in-situ matchup database that was used to perform the round-robin exercise and designate the best performing algorithm per optical waterclass.

The improved performance in the final product through blending should reduce the class uncertainties, therefore reducing the final per-pixel uncertainty estimates. Additionally, the ability to capture some of the structure in the error distribution across water classes (for example, some chl algorithms tend to

underestimate at high chl and overestimate at low chl concentrations) allows the water class approach to have a better representation of per pixel uncertainty than estimates using all in-situ data to give a single uncertainty estimate.

Figure 7 shows that the product blending has a net improvement on the product uncertainty. The effect is not uniform in time or space as the abundance of different water classes varies through the year and in some cases the blended product has a greater RMSD than the non-blended product, though this is certainly not the dominant situation.

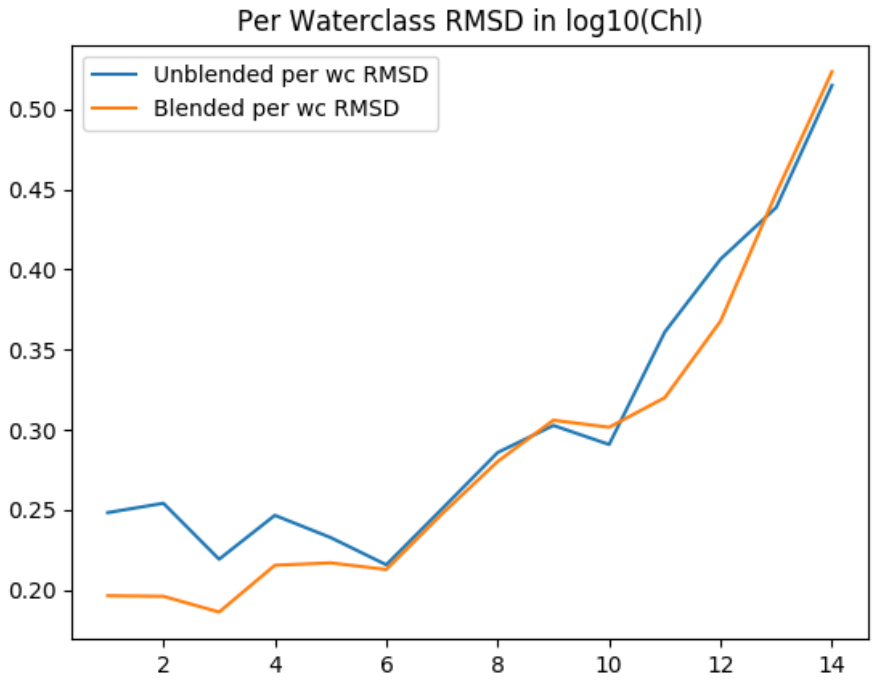


*Figure 7: Blended Chlorophyll-a product has better performance (lower RMSD) in almost all water classes and especially in those dominating the largest area of the global oceans (water classes 1-5)*

### 4.5.2 Limitations and requirements for in-situ data

As this method relies on knowledge of algorithm performance across a range of optical conditions there is a strong dependence on a large, high quality in-situ database for matching against the satellite data record. This requirement is further emphasised if the waterclasses and blending is also used to assign uncertainty estimates. There is therefore an ongoing requirement for in-situ data collection across a range of optical water types in order to improve the per waterclass performance characterisation, especially as new sensors are launched and new algorithms developed