ECMWF COPERNICUS REPORT

Copernicus Climate Change Service

Product Quality Assessment Report

Ocean Colour Version 6.0

Issued by: PML Applications / Thomas Jackson

Date: 22/03/2023

Ref: WP2-FDDP-2022-04_C3S2-Lot3_ PQAR-of-v6.0-OceanColour-product_v1.1

Official reference number service contract:

2021/C3S2_312a_Lot3_METNorway/SC1

PROGRAMME OF THE EUROPEAN UNION

IMPLEMENTED BY ECMWF

Please cite this data product as: Jackson et al. (2022) *Ocean colour daily data from 1997 to present derived from satellite observations, Version 6.0*. Copernicus Climate Change Service (C3S) Data Store (CDS), accessed <DD Month YYYY>, https://doi.org/10.24381/cds.f85b319d.

To cite this document, please use: Jackson, T., Calton, B., Hockley, K. (2023) C3S Ocean Colour Version 6.0: Product Quality Assessment Report. Issue 1.1. E.U. Copernicus Climate Change Service. Document ref. WP2-FDDP-2022-04_C3S2-Lot3_ PQAR-of-v6.0-OceanColour-product.

Contributors

PML APPLICATIONS
Thomas Jackson
Ben Calton
Kim Hockley

History of modifications

Version	Date	Description of modification	Chapters / Sections
1.0	06/05/2022	Creation. *Jackson et al.* (2020) C3S Ocean Colour Version 5.0: Product Quality Assessment Report, issue 1.2. E.U. Copernicus Climate Change Service, document ref. D2.OC.2-v5.0_PQAR_of_v5.0_Ocean_Colour_product_v1.2 was used a basis for this document.	All
1.1	22/03/2023	Document revised in response to feedback from independent review, and finalised for publication.	All

List of datasets covered by this document

Deliverable ID	Product title	Product type (CDR, ICDR)	Version number	Delivery date
WP2-FDDP-2022-04_ C3S2-Lot3_DatasetDelivery-of-v6.0-OC-CDR_v1.0	Ocean Colour ECVs	CDR	6.0	15 June 2022
WP2-ICDR-OC-v6.0_C3S2-Lot3_DatasetDelivery-of-v6.0-OC-ICDR_v1.0	Ocean Colour ECVs	ICDR	6.0	27 July 2022
WP2-ICDR-OC-v6.0_C3S2-Lot3_DatasetDelivery-of-v6.0-OC-ICDR_v1.0	Ocean Colour ECVs	ICDR	6.0	31 October 2022
WP2-ICDR-OC-v6.0_C3S2-Lot3_DatasetDelivery-of-v6.0-OC-ICDR_v1.0	Ocean Colour ECVs	ICDR	6.0	31 January 2022

Related documents

Reference ID	Document
C3S2_ATBD	Jackson, T., Hockley, K., Calton, B., Chuprin, A. (2022) *C3S Ocean Colour Version 6.0: Algorithm Theoretical Basis Document*. Issue 1.1. E.U. Copernicus Climate Change Service. Document ref. WP2-FDDP-2022-04_C3S2-Lot3_ATBD-of-v6.0-OceanColour-product.
C3S2_PQAD	Jackson, T., et al. (2023) *C3S Ocean Colour Version 6.0: Product Quality Assurance Document*. Issue 1.1. E.U. Copernicus Climate Change Service. Document ref. WP1-PDDP-OC-v6.0_C3S2-Lot3_PQAD-of-v6.0-OceanColour-product.
C3S-SQAD	Jackson, T. (2023) *C3S Ocean Colour Version 6.0: System Quality Assurance Document*. Issue 1.1. E.U. Copernicus Climate Change Service. Document ref. WP3-SQAD-OC-v6.0_C3S2-Lot3_SQAD-of-v6.0-OceanColor-product.
C3S2_PUGS	Jackson, T., Calton, B., Hockley, K. (2023) *C3S Ocean Colour Version 6.0: Product User Guide and Specification*. Issue 1.1. E.U. Copernicus Climate Change Service. Document ref. WP2-FDDP-2022-04_C3S2-Lot3_PUGS-of-v6.0-OceanColour-product.
OC_CCI-PVASR	OC-CCI Product Validation and Algorithm Selection Reports (atmospheric and in water reports all available). Available at https://climate.esa.int/en/projects/ocean-colour/key-documents/ [last accessed 8th March 2023].
OC_CCI-PVIR	OC-CCI Product Validation and Inter-Comparison Report. Available at https://climate.esa.int/en/projects/ocean-colour/key-documents/ [last accessed 8th March 2023].
OC_CCI-CECR	OC-CCI Comprehensive Error Characterisation Report. Available at https://climate.esa.int/en/projects/ocean-colour/key-documents/ [last accessed 8th March 2023].

Acronyms

Acronym	Definition
AERONET-OC	Aerosol Robotic Network - Ocean Colour
ATBD	Algorithm Theoretical Basis Document
BOUSSOLE	Buoy for the acquisition of a long-term optical time series
C3S	Copernicus Climate Change Service
CDR	Climate Data Record
CECR	Comprehensive Error Characterisation Report
Chl	Chlorophyll
ECV	Essential Climate Variable
ESA	European Space Agency
GAC	Global Area Coverage
GCOS	Global Climate Observing System
HPLC	High Performance Liquid Chromatography
ICDR	Interim Climate Data Record
LAC	Local Area Coverage
MERIS	Medium Resolution Imaging Spectrometer
MERMAID	MEris MAtchup In-situ Database
MLAC	Merged LAC
MOBY	Martine Optical Buoy
MODIS	NASA Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
NOMAD	NASA bio-Optical Marine Algorithm Dataset
OC	Ocean Colour
OC-CCI	Ocean Colour Climate Change Initiative
OLCI	Ocean and Land Colour Instrument
PQAD	Product Quality Assurance Document
PQAR	Product Quality Assessment Report
PUGS	Product User Guide and Specification
PVASR	Product Validation and Algorithm Selection Report
PVIR	Product Validation and Inter-comparison Report
RMSD	Root- mean-square difference
RMSE	Root-mean-square error
Rrs	Remote Sensing Reflectance
SeaBASS	SeaWiFS Bio-optical Archive and Storage System
SeaWiFS	Sea-Viewing Wide Field-of-View Sensor
SPG	South Pacific Gyre
SQAD	System Quality Assurance Document
VIIRS	Visible Infrared Imaging Radiometer Suite

General definitions

Atmospheric Correction

Atmospheric correction is the process of removing the effects of the [atmosphere](https://en.wikipedia.org/wiki/Atmosphere) on the satellite images so that we can obtain information about the surface of the Earth. Atmospheric effects in optical remote sensing are significant and complex but can be largely considered as absorbing or scattering factors. Also keep in mind that the light reaching the satellite borne sensor has passed through the atmosphere twice, from the sun to the surface and then back to the sensor. For ocean colour remote sensing, the surface signal from the ocean is typically ≤10% of the total signal received by the satellite (with the other 90+% coming from the atmosphere).

Binning

In the context of this document, binning refers to the process of aggregating data into bins. This is an essential process when it comes to merging data from multiple sensors. Each remote sensing platform used to observe the Earth collects data in a manner that is constrained by the sensor design and satellite orbit. This means that different sensors will collect data at difference spatial resolutions and viewing geometries. Once the data has been processed at its native resolution it can be binned onto a defined grid for easier use.

Chlorophyll-a

Chlorophyll-a is a green pigment and the most prevalent photosynthetic pigment in both terrestrial and marine photosynthetic organisms. As an indicator of phytoplankton abundance, and therefore the base of the marine foodweb, chlorophyll-a concentration (chl-a) is recognised as an Essential Climate Variable. Oceanic chlorophyll-a is usually measured in units of mg m^{-3}, with concentrations ranging over multiple orders of magnitude.

Climate Data Record

The term Climate Data Record has a specific definition developed by the CEOS-CGMS Joint Working Group on Climate in 2020. The CEOS definition scheme defines three types of climate data records: 1) Fundamental Climate Data Records (FCDRs) consist of a consistently processed time series of uncertainty-quantified sensor observations calibrated to physical units, located in time and space, and of sufficient length and quality to be useful for climate science or applications; 2) Climate Data Records (CDRs) consist of a consistently processed time series of uncertainty-quantified retrieved values of a geophysical variable or related indicator, located in time and space, and of sufficient length and quality to be useful for climate science or applications; 3) Interim Climate Data Records (ICDRs) are consistently processed times series of uncertainty-quantified estimates of CDR values produced with better timeliness than, but otherwise minimising differences with, the estimated CDR values.

Essential Climate Variable

An Essential Climate variable (ECV) is a physical, chemical, or biological variable or a group of linked variables that critically contributes to the characterisation of Earth's climate. The Global Observing

Systems Information Center (GOSIC) provides further background, definitions, requirements, network information, and data sources for the ECVs.

In-water algorithms
The term ‘in-water’ is used to describe algorithms that estimate in-water properties of the surface water from the surface reflectance signal. These algorithms often provide estimates of concentrations of substance (such as chlorophyll-a or sediment) but could also provide estimates of Inherent Optical Properties (IOPs) such as absorption or scattering.

Level-[x] remote sensing data
Within remote sensing and ocean colour applications, datasets are often described in terms of levels. The level is representative of the amount of processing that has been performed. **Level-0** is the rawest data format available. It is full resolution data, as it comes from the instrument, with some processing applied to remove artefacts from data communication between the satellite and the ground stations. **Level-1** data is full resolution sensor data with time-referencing, ancillary information, including radiometric and geometric calibration coefficients and georeferencing parameters, computed and added to the file. **Level-2** refers to derived geophysical variables (such as water-leaving reflectance or ocean colour products) at full resolution. This will have required processing to remove the atmospheric component of the signal. Pixels will also be masked by use of data quality flags. **Level-3** data is a binned version of the level-2 products at a given temporal and spatial resolution.

Masking
Masking is the process of setting pixels to NaN or blank values where a flag has been raised that the data would not be of sufficient quality for the intended purpose (or processing has failed). There are many factors that can lead to remote sensing data being of insufficient quality for a climate data record, so we shall not list them all here, but commonly applied masks in ocean colour remote sensing include cloud, cloud shadow, land, glint, and algorithm failure masks.

Match-ups
In the context of this document a match-up refers to a matched pair of in situ and remote sensing data. These measurements are matched based on their time and location information where some permitted time or space offset is permitted, for example we might match an in situ measurement to the closest pixel on the satellite data grid for the same day of observation. It is also of note that these measurements are also made using information at very different scales; an in situ measurement of chlorophyll-a might be from a litre of filtered seawater while the remote sensing estimate may be derived over a pixel 1km square (or larger).

Ocean Colour
When sunlight passes through the atmosphere and enters the ocean, the different colours in the light spectrum are absorbed and scattered as they encounter different particles and substances on their journey. This absorption and scattering of the visible spectrum confers a spectral signal to the reflected light measured by the sensors as ‘ocean colour’ data.

The Global Climate Observing System (GCOS) define ocean colour as:

> “Ocean colour is the radiance emanating from the ocean normalized by the irradiance illuminating the ocean. Products derived from ocean colour remote sensing (OCRS) contain information on the ocean albedo and information on the constituents of the seawater, in particular, phytoplankton pigments such as chlorophyll-a. OCRS products are used to assess ocean ecosystem health and productivity, and the role of the oceans in the global carbon cycle, to manage living marine resources, and to quantify the impacts of climate variability and change.”

Phytoplankton

Phytoplankton are aquatic microscopic photosynthetic organisms. This group includes some bacteria, protists, and single-celled plants. There is a great diversity in appearance and function across phytoplankton, with orders of magnitude in size between the largest and smallest phytoplankton. Given their photosynthetic abilities, phytoplankton form the base of the marine food web. Phytoplankton growth primarily depends on the availability of sunlight and nutrients.

Remote sensing reflectance

Remote Sensing Reflectance, $R_{rs}(\lambda)$, has units of sr^{-1} (per steradian) and is the water-leaving radiance, corrected for bidirectional effects of the air-sea interface and sub-surface light field, normalised by downwelling solar irradiance, $Ed(\lambda)$, just above the sea surface. This is usually measured at multiple wavelengths by a given ocean colour sensor, so R_{rs} is a spectral product. R_{rs} is the primary variable of the Ocean Colour ECV, and the chlorophyll-a products are derived from the Rrs data.

Table of Contents

Scope of the document

This document summarises the results of the validation of Version 6.0 of the C3S Ocean Colour products. Given that the C3S OC products are based on a processing chain originally developed by the ESA OC-CCI project, the reader is referred to OC-CCI documentation where applicable.

Executive summary

This document provides details of the quantitative and qualitative assessment of the chlorophyll and remote sensing reflectance (Rrs) data, provided through the C3S Ocean Colour dataset version 6. This assessment follows the assessment scheme and methods described in the associated Product Quality Assurance Document (C3S2_PQAD). The primary input data used are the OC-CCI in-situ database v8.0, which are matched against the output products for quantitative product assessment.

The in-situ database contains many tens of thousands of chlorophyll-a and Rrs measurements and provides > 19,000 matchups against satellite data for assessment of chlorophyll-a and Rrs. Analysis of matchups against the in-situ database show that the chlorophyll product shows a very small negative bias, a small root-mean-square error (RMSE), and strong correlation coefficient. The slope of the regression is less than one, which is mainly forced by an underestimate at the higher chlorophyll-a concentrations.

In comparison with previous versions of the product the v6.0 shows an improved performance and also has enhanced coverage from the inclusion of OLCI 3B into the processing chain.

Both the Rrs and chlorophyll-a products provided appear to be in-line with the requirements set out by the user community and GCOS in terms of resolution, accuracy and stability. A more rigorous metrological assessment of product accuracy and stability would be of use and this remains an ongoing goal of the OC-CCI project.

This document first presents a brief overview of the in-situ datasets input into the quality assessment activity. The quality checks post-binning are then outlined, before the results from the comparison to in-situ data are presented. Finally, the document clarifies how the product aligns with the user requirements of the C3S service, in light of the quality assessment.

1. Product validation methodology

This document provides a summary of the validation of the Ocean Colour products (V6.0) as provided through the C3S2 project. The Ocean Colour product set consist of remote-sensing reflectance at multiple wavelengths and chlorophyll-a concentration. To maximise product quality and accelerate the cycle of processor improvement, multiple inspection/validation exercises are performed on intermediate processing outputs, prior to the generation of the final products. This includes checks on the input data streams and pre- and post- bias correction checks. The final products (produced through implementation of the optimal algorithms designated in the Algorithm Theoretical Basis Document; C3S2_ATBD) are then compared with in-situ validation data. In-situ matchups (same day of measurement) are used to generate multiple metrics of performance (correlation coefficient, slope and intercept of type-2 linear regression, bias, root mean square difference) and longer time-series are used to check the stability of the final record against in-situ measurements (even though direct matchups may not be available). Finally, the product performance is compared against the requirements set out by GCOS for climate data records of Ocean Colour. Further, more detailed information on the validation approach can be found in the Ocean Colour, Version 6.0 Product Quality Assurance document (C3S2_PQAD).

2. Validation results

2.1 Input dataset Quality Control

This PQAR applies to C3S Ocean Colour products as released in May 2022. The document provides a validation and quality assessment of the products, from input data sources to final products, using in situ data and similar remote-sensing data products. The products covered are:

- Chlorophyll-a
- Remote Sensing Reflectance (Rrs) at the following wavelengths (nm):
  - 412
  - 443
  - 490
  - 510
  - 555
  - 670

See the Product User Guide and Specification (C3S2_PUGS) for more information on these various data layers.

2.1.1 In-situ data coverage

The reference dataset used to validate the C3S2 OC product is the OC-CCI in-situ database v8.0. This covers 1997-2021 and has a global distribution. It includes data from the MOBY[1], BOUSSOLE[2], AERONET-OC[3], SeaBASS[4], NOMAD[5], MERMAID[6], CoastColour[7], AWI-Polarstern cruises[8], and the Tara expedition[9].

This merged database contains 57,811 observations of “Rrs” and combined 85,752 “chla” (either “chla_fluor” or “chla_hplc” for a sample). The spatial distribution of the chlorophyll values for the combined dataset (Figure 1) shows a good agreement with known biogeographical features, such as low chlorophyll values in the subtropical gyres and high values in temperate, coastal and upwelling regions. Many regions have a good spatial coverage (e.g. Atlantic and Pacific Oceans), while others are less well sampled (e.g. Southern and Indian Oceans).

[1] https://mlml.sjsu.edu/moby/ [Last accessed 9th March, 2023]
[2] http://www.obs-vlfr.fr/Boussole/html/home/home.php [Last accessed 9th March, 2023]
[3] https://aeronet.gsfc.nasa.gov/new_web/ocean_color.html [Last accessed 9th March, 2023]
[4] https://seabass.gsfc.nasa.gov/wiki/System_Description [Last accessed 9th March, 2023]
[5] https://seabass.gsfc.nasa.gov/wiki/NOMAD [Last accessed 9th March, 2023]
[6] http://mermaid.acri.fr/home/home.php [Last accessed 9th March, 2023]
[7] https://www.coastcolour.org/db_in-situ.html [Last accessed 9th March, 2023]
[8] https://www.awi.de/en/expedition/research-vessel-and-cutter/polarstern.html [Last accessed 9th March, 2023]
[9] https://fondationtaraocean.org/en/expedition/tara-oceans/ [Last accessed 9th March, 2023]

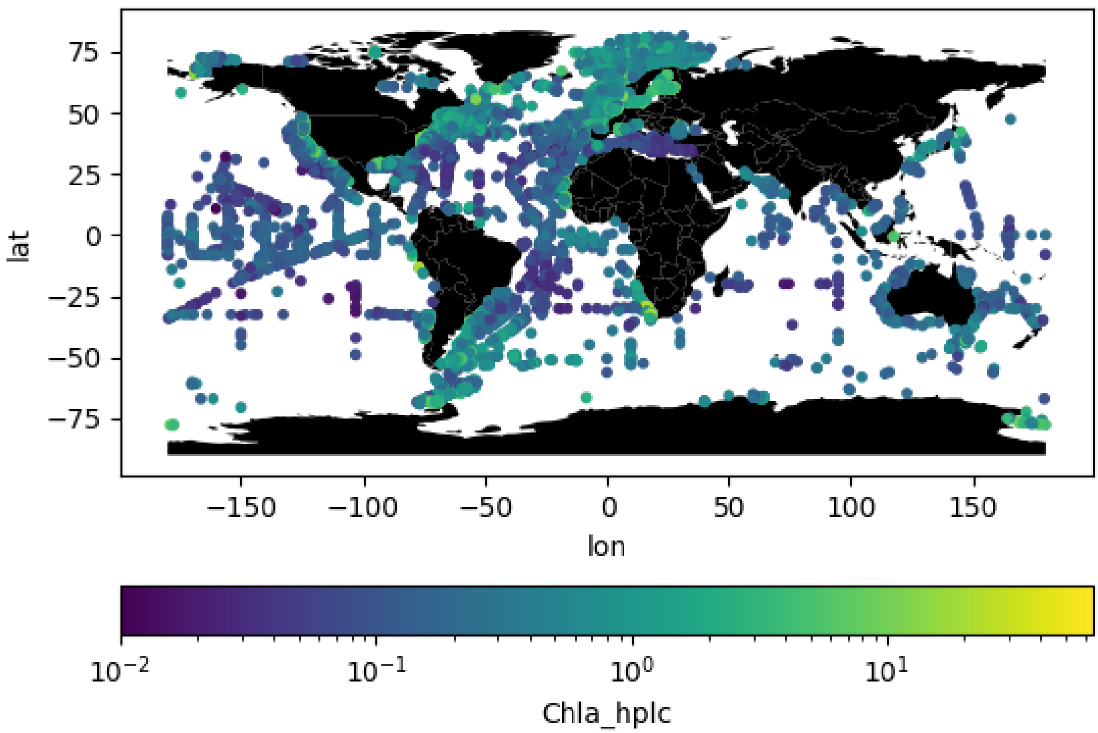

Figure 1: Global distribution of chlorophyll-a concentration coloured based on the observed value. All chlorophyll data were considered, both High Performance Liquid Chromatography (HPLC) and extracted fluorometric data, but for a given station data HPLC were selected in preference to extracted pigment fluorometry data, if available. Lower values are seen in the central ocean gyres while coastal waters and temperate latitudes tend to support higher peak chlorophyll-a concentrations.

2.1.2 Input Satellite Data availability

Following the structure of the PQAD, the first consideration for product quality is the input data sources. The currently released ECV dataset contains products in the time range 1997-2022 and **Figure 2** shows the number of contributing granules per sensor over this period. A granule is a tile of data from the satellite (the data is broken into tiles for transmission to receiving stations). It is of note that MODIS and VIIRS have a higher number of granules as the files each cover a smaller region of the earth (OLCI, MERIS and SeaWiFS are mostly 'half-orbit' input files).

The dips (Figure 2) in the number of contributing granules have been checked against data records of the relevant sensor operating agencies and are confirmed as either 1) dates for which standard daily coverage is not available/provided or 2) dates for which the granules are 'blacklisted' from processing due to known issues with the data. This lack of coverage or blacklisting can be due to a number of reasons such as sensor maintenance, satellite orbit corrections, data corruption etc. It is worth noting that MODIS and Suomi-VIIRS data are available for the period 2020-present but they have been excluded from the CDR merged record due to concerns about their quality in the most recent years, as discussed in section 1.2.

Due to the different sensor designs and viewing geometries, the granules from each of the contributing sensors provide differing numbers of pixels to the merged record. The daily number of pixels from each sensor that contribute to the binned products is shown in **Figure 3**. Here we can see that technological improvements in remote sensing technologies over the last 25 years have allowed the number of observations to increase considerably over the course of the record.

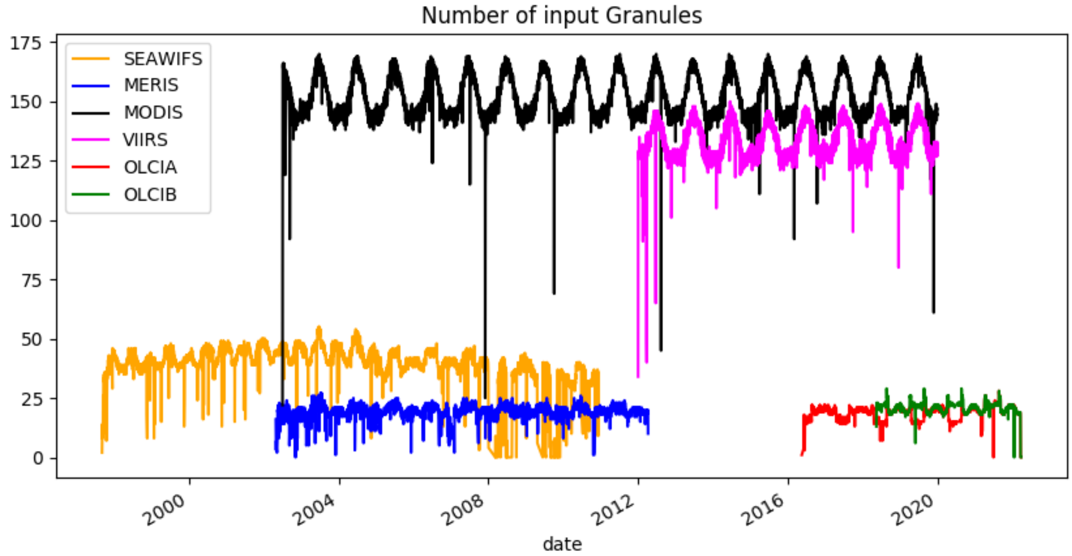

Figure 2: Daily number of granules from each contributing sensor to the CDR. Note that 'granules' for OLCI-A, OLCI-B, and MERIS are ingested as half-orbits (composites of multiple smaller granules); the SeaWiFS input includes both Global Area Coverage (GAC) half orbits and Merged Local Area Collection (MLAC) granules; MODIS and VIIRS granules are only ingested as separate granules (no half-orbits).

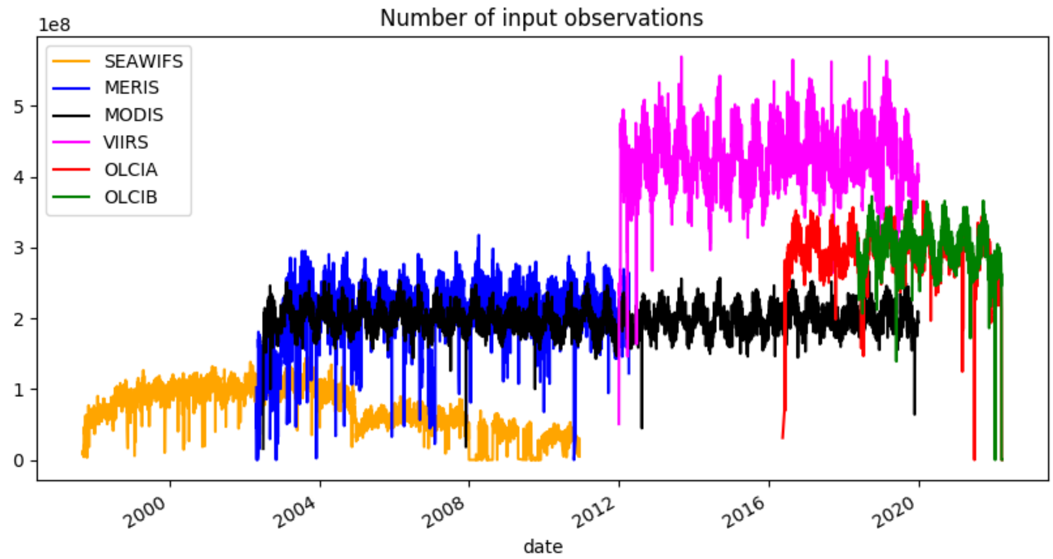

Figure 3: Daily number of pixels contributing to the merged record from each of the sensors contributing to the C3S Ocean Colour products.

2.2 Processing Checks

2.2.1 Post binning and bandshifting data

The first stage of the CDR processing chain is to bin all the input sensors onto the same spatial grid at a daily temporal resolution. As noted in the PQAD, a number of regions are examined at the intermediate processing stages, with the central South Pacific Gyre (SPG) used as a region that exhibits a relatively small annual cycle and is highly stable (SPG defined here as Lat: 20°S to 30°S and Lon: 110°W to 120°W). The stability of these regions makes it easier to identify significant anomalies that might need further investigation in-case of issues in the processing chain or atmospheric correction.

We then bandshift the data globally and create inter-sensor bias maps to correct for inter-sensory differences when observing the same regions of the ocean. As described in the PQAD, these bias maps (such as that shown in **Figure 4**) are visually inspected for artifacts or anomalies.

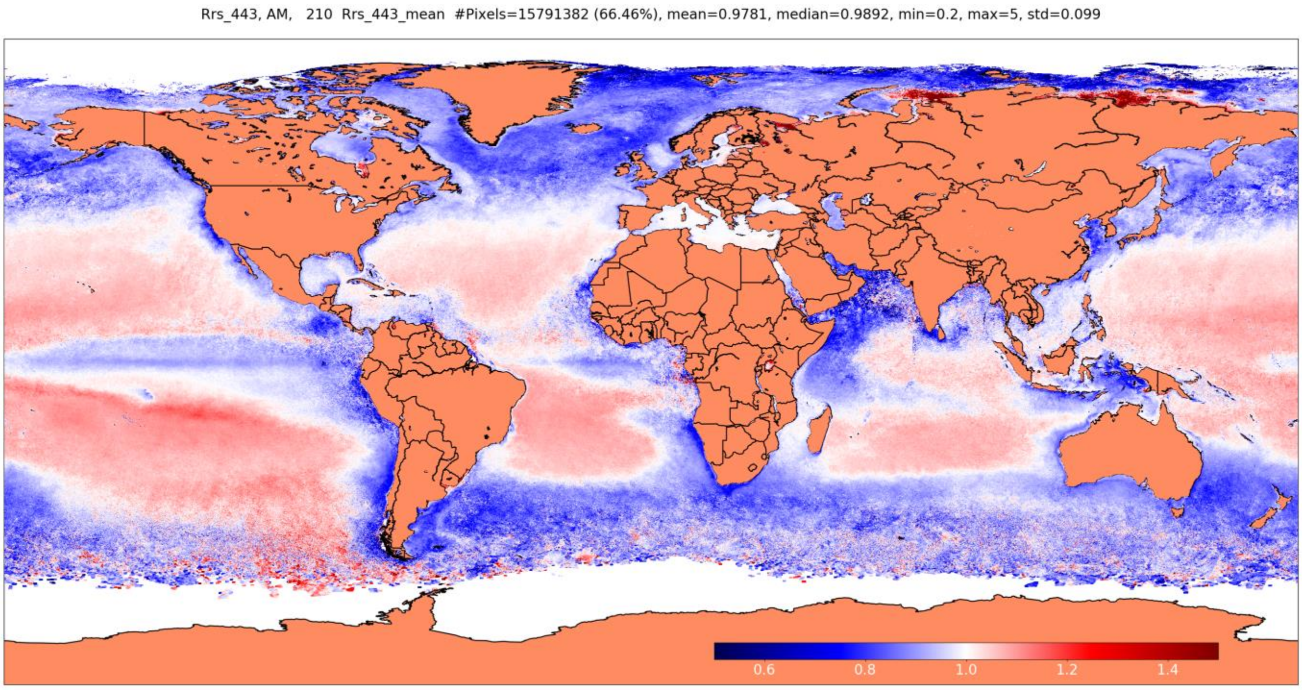

Figure 4: Inter sensor bias correction map for Julian day 210 and band 443nm between MODIS and MERIS sensors. The colour scale shows values where the ratio between MODIS and MERIS estimates is >1 (red) or <1 (blue). This map shows that there is a tendency for one sensor to estimate a higher 443 reflectance in the clear gyre waters, but we do not observe any significant artefacts/anomalies.

Following visual inspection and expert acceptance of the bias maps, we then proceeded with the de-biasing of the single sensor data streams. Regional time-series for the South Pacific Gyre (at 443nm) per sensor are shown in **Figure 5** and **Figure 6**.

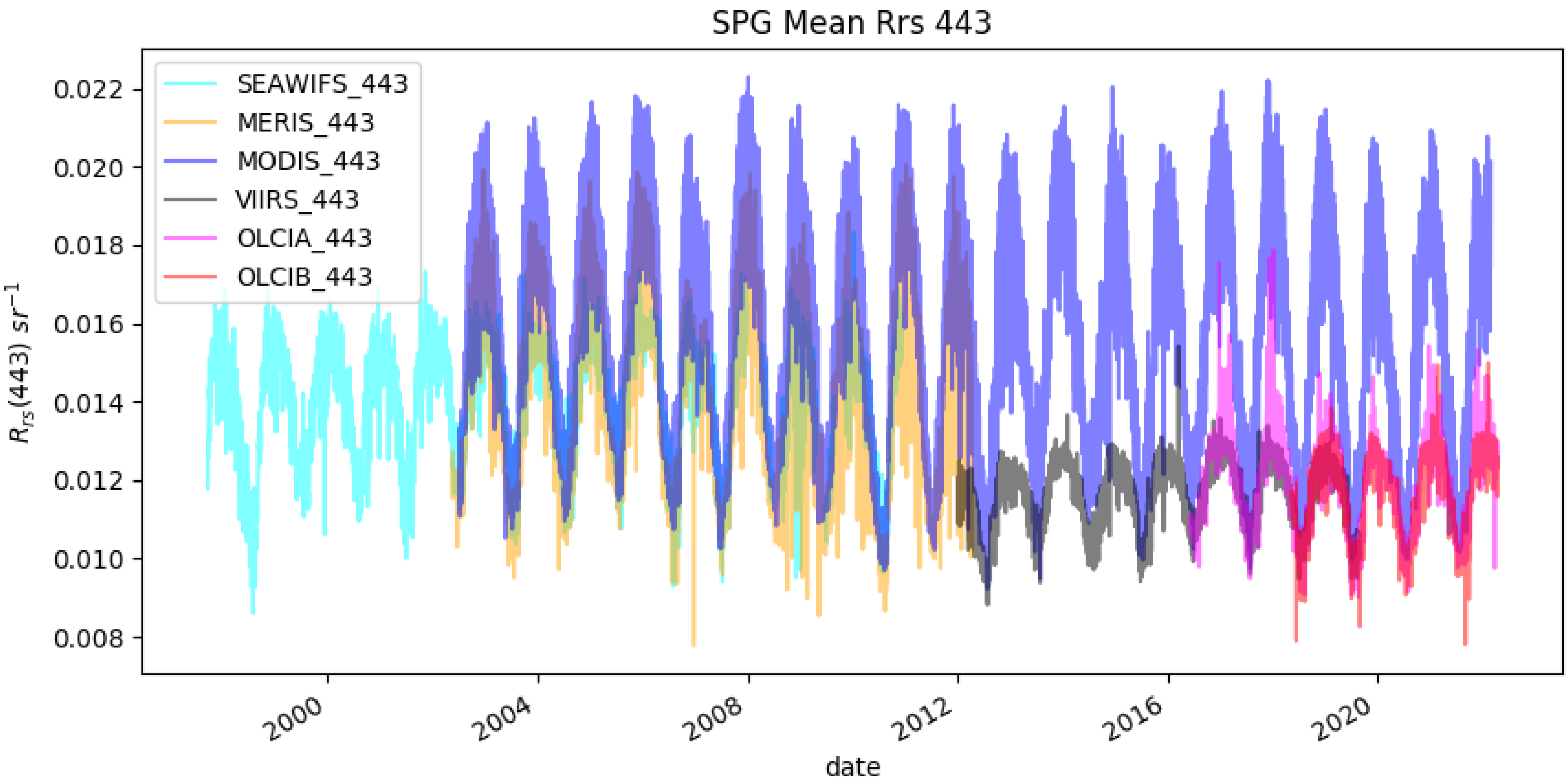

Figure 5: Overlain timeseries of mean Rrs (443 nm) values for a 10 degree square region in the South Pacific Gyre for each of the contributing level2 data streams. Note that transparency is used to emphasize where there is overlap in mean Rrs derived from different sensors.

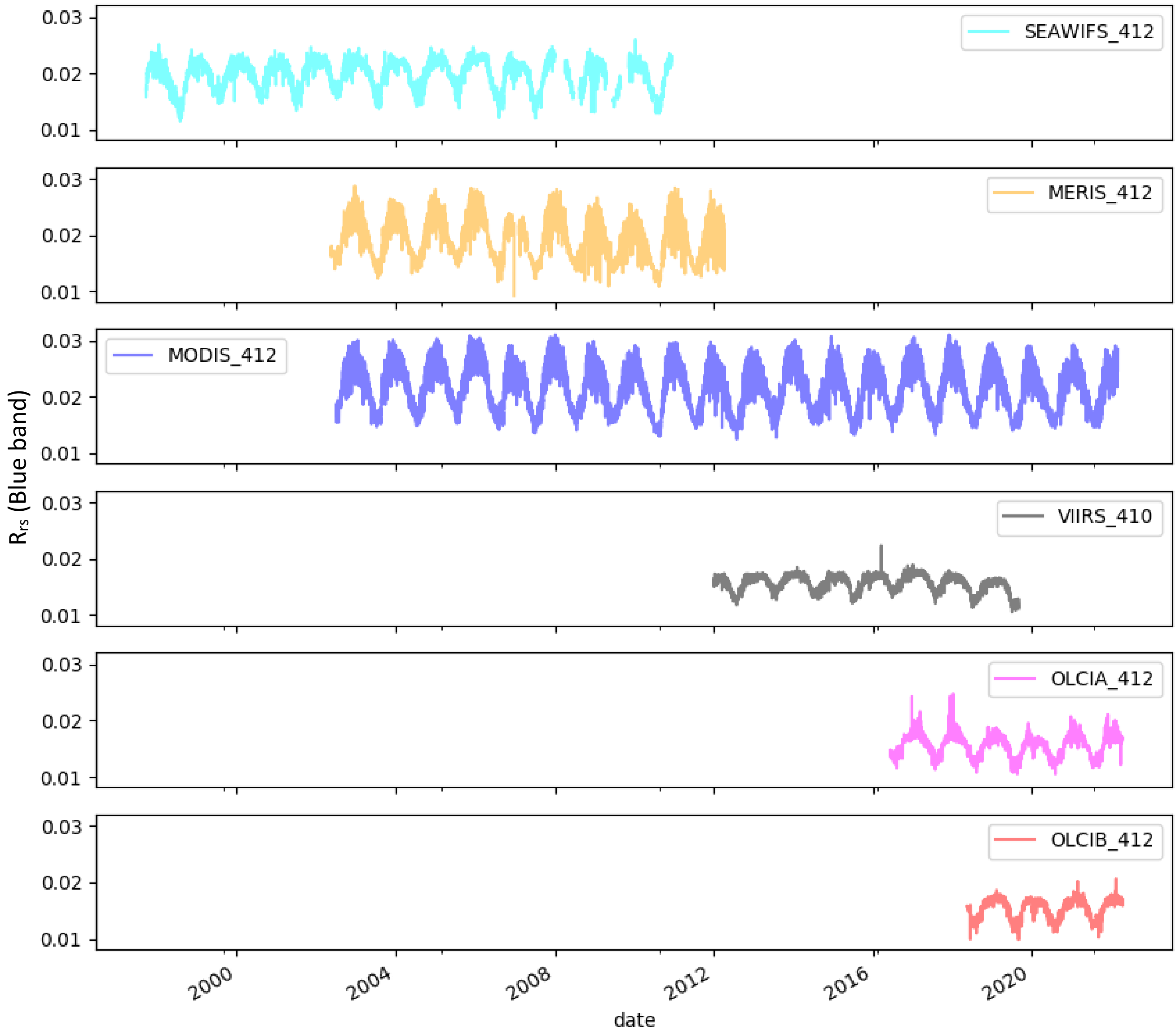

Figure 6: Individual plots of timeseries of mean depending on sensor) values for a 10 degree square region in the South Pacific Gyre for each of the contributing level2 data streams.

Although the difference between sensors might be considered alarming for climate data, the difference between sensors observing the same region of the ocean at the same time simply highlights the need for inter-sensor bias-correction. All sensors agree in terms of the seasonality of the Rrs signal for the region, but the differing dynamic ranges between sensors is notable

The bias correction process is designed to deal with this issue. The primary concern to check for at this stage (post L3 binning) is that there are no obvious trends that are seen in one sensor compared to the other sensors. The latter portion of the VIIRS record (2020-2022) is not used because of this (it was identified at this stage), and the ingested VIIRS record was therefore truncated to the end of 2019.

2.2.2 Post bias correction data

The post bias correction checks for Version 6.0 appear to confirm that the bias correction scheme is not introducing artifacts to the record, and is doing a good job of harmonising the observations from differing sensors to the record. This is clearly demonstrated by comparing **Figure 5** with **Error! Reference source not found.**. It is worth noting that the bias correction scheme should not remove or add trends within a sensor record as the bias correction is calculated and applied per Julian day (meaning that the bias correction scaling for a given pixel is the same for a given day of the year).

Although it is not merged into the final record, we have processed the data for the VIIRS and MODIS sensors to the end of the record to verify it was the correct decision. The post bias correction data in the Mediterranean clearly shows that the VIIRS sensor became less reliable from 2020 onwards. This is shown in **Figure 8**. It is worth noting that the increased noise and range of values for VIIRS from 2020 onwards in the Mediterranean region must be related to the input data, because the bandshifting and bias correction schemes do not change from year to year.

2.2.3 Output file attribute checks

The product standardiser process was run over all output product files to ensure metadata attributes were not malformed. Individual products were checked for corruption within each file, using tests of variable size, value ranges, and testing file accessibility/readability by a NetCDF interpreter.

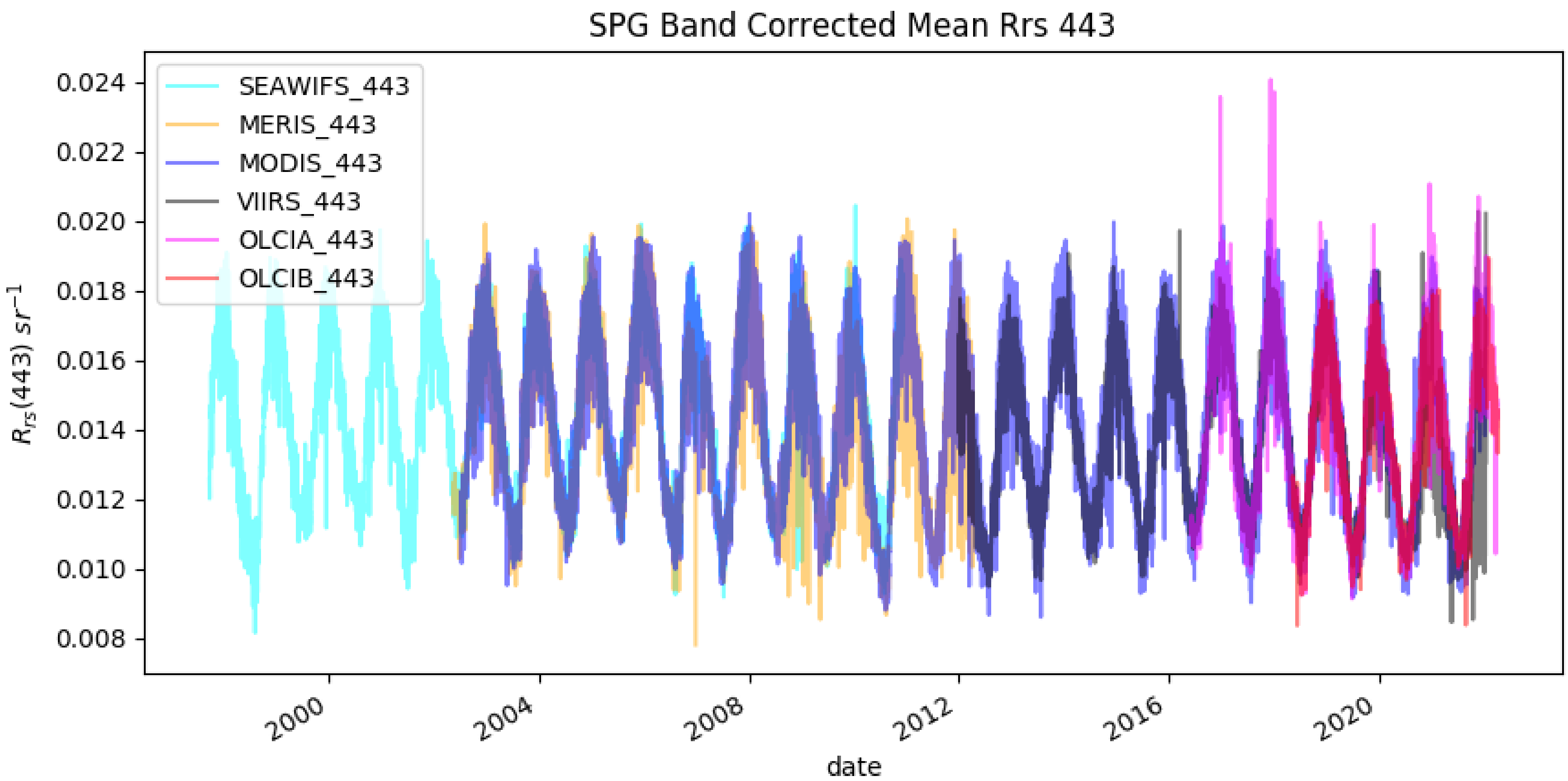

Figure 7: Post bias correction SPG average Rrs (443 nm) time series per-sensor. Strong agreement now seen between all the contributing sensors. Note that transparency is used to emphasize where there is overlap in mean Rrs derived from different sensors.

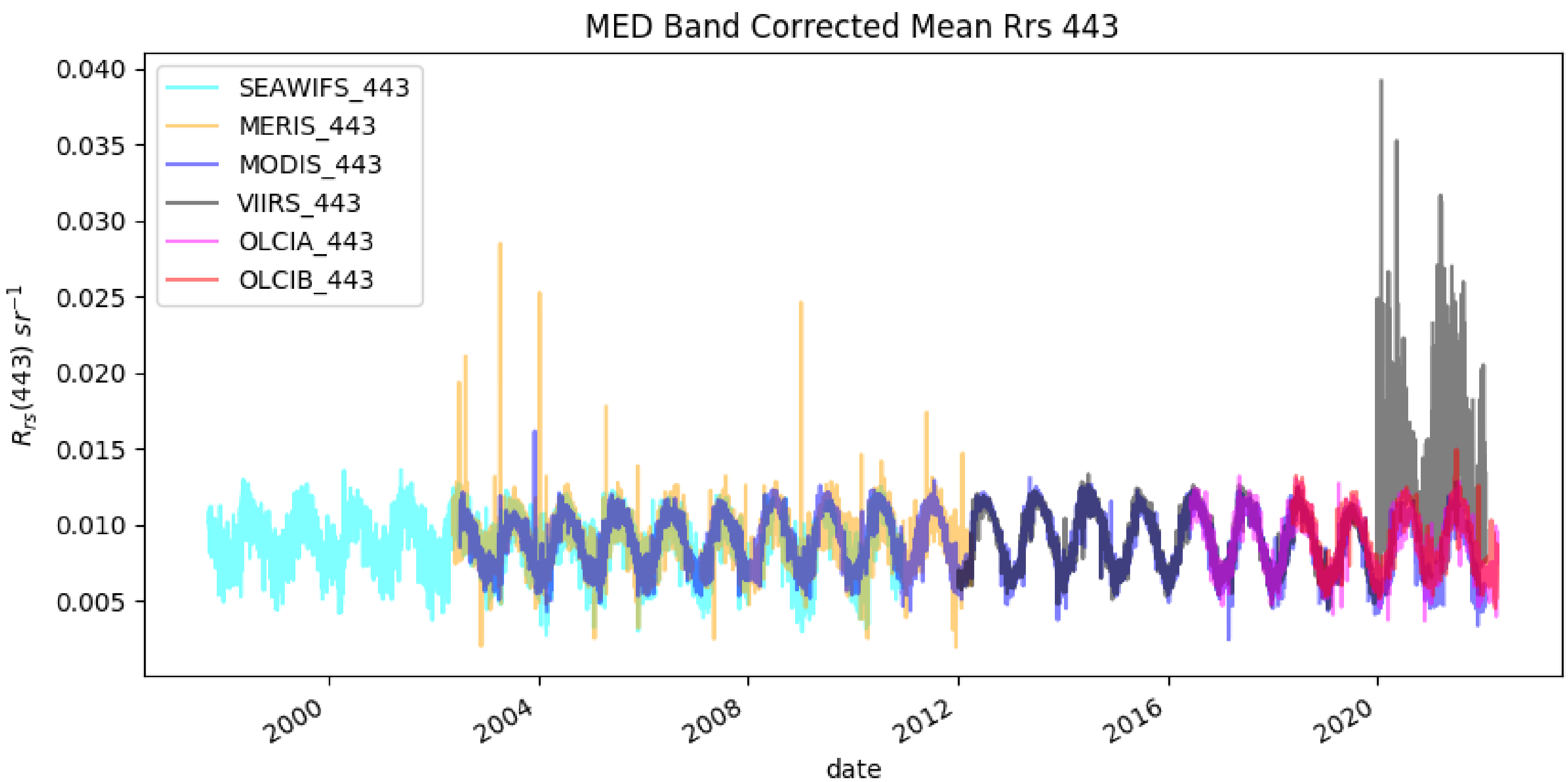

Figure 8: Post bias correction time-series, per-sensor, for a 10 degree by 10 degree region in the Mediterranean Sea. Note the massive increase in noise and dynamic range in the VIIRS data from 2020 onwards. This data is not used in the final merged record but we have plotted it here for demonstration purposes. Note that transparency is used to emphasize where there is overlap in mean Rrs derived from different sensors.

2.3 Comparison to in-situ datasets

2.3.1 Overall performance

From the complete set of match-up data, the chlorophyll product shows a very small bias, a small root-mean-square error (RMSE), and strong correlation coefficient (**Figure 9**). The slope of the regression is less than one, which is forced by an underestimate at higher concentrations. The slope increases towards unity if higher concentrations are removed from the analysis (for data with in-situ chl-a concentrations of 1.0 or below the slope of regression is 0.889). The statistics also show improved performance if only HPLC data are used (**Figure 9** shows merged HPLC and extracted fluorescence in situ data).

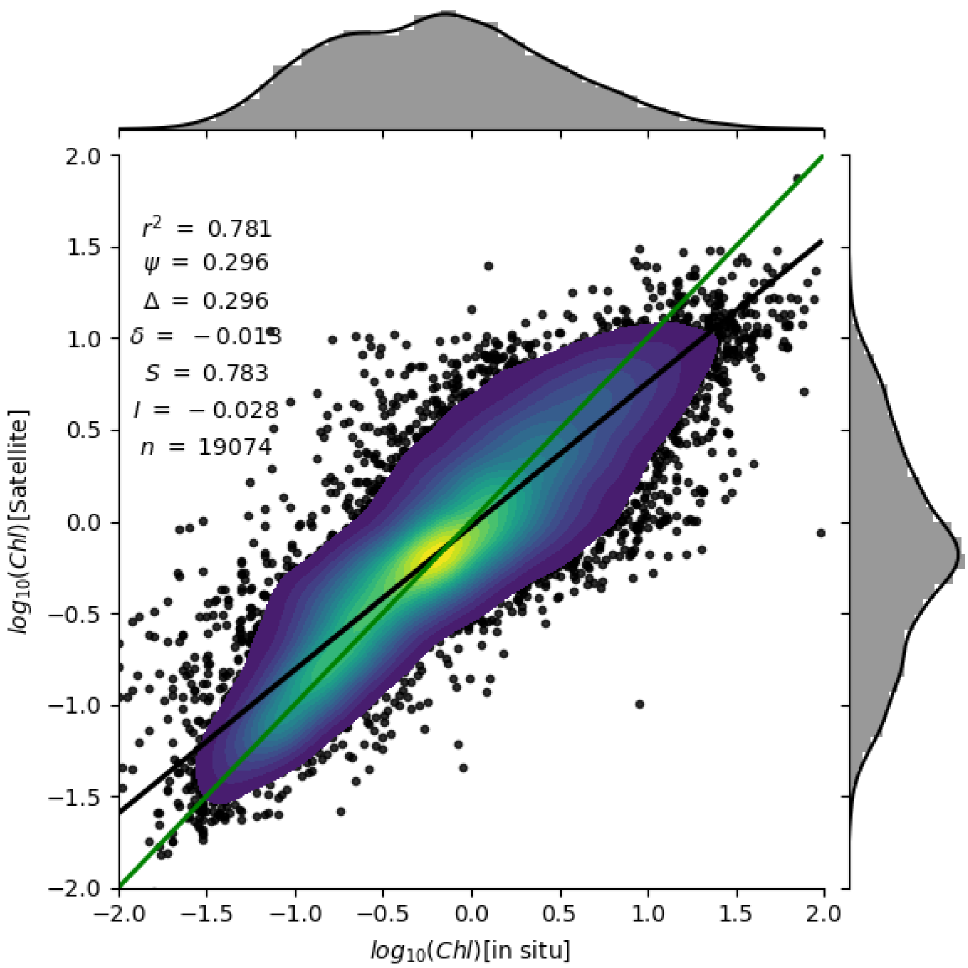

Figure 9: OC_CCI v6 chlorophyll-a match-up density plot. The 1:1 line(green), linear regression (black), squared correlation coefficient (r2), RMSE (ψ), unbiased RMSE (Δ), bias (δ), Slope (S) and intercept (I) of regression, and number of match-ups (n) are shown. The density colour scale runs from blue (low) to yellow (high). The grey plots on the top and right of the figure are histograms along the x and y axes.

2.3.2 Per water class breakdown

The OC-CCI project classified waters into optical water types in order to utilise optimal in-water algorithms for product retrieval (see OC-CCI-CECR and C3S2_ATBD for details). This classification scheme also enabled an assessment of product performance to be implemented for each of the optical water types. While full details are given in the [OC_CCI-CECR] document, the figures below summarise the performance of the ocean colour products (in terms of bias and RMSE) for each of the 14 optical water types. Water class 1 is representative of the most oligotrophic central gyres and classes 12-14 represent more turbid, typically coastal waters. The transition between these extremes is covered by the sequence of water classes from 2-11. For Rrs products, one also has to consider the variability in performance with both optical water type and wavelength.

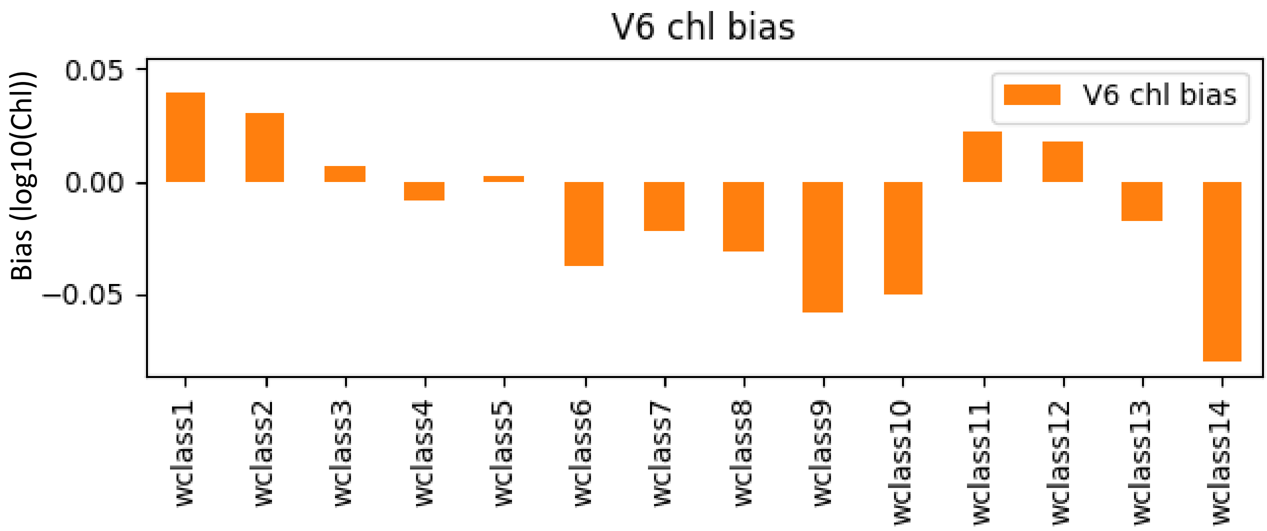

Figure 10: Chlorophyll-a average bias (log10(chl)) per optical water class (wclass).

Given that for chl-a the slope of regression is less than one, we would expect to see a transition from overestimates in clear waters to underestimates in more turbid waters. When the water classes are roughly ordered in order of increasing chlorophyll-a concentration (as done in **Figure 10**), we do indeed observe that classes (1-3) seem to exhibit a slight positive bias (overestimate) in chlorophyll-a estimates while the higher numbered water classes generally show a slightly negative bias (underestimate) for chlorophyll-a.

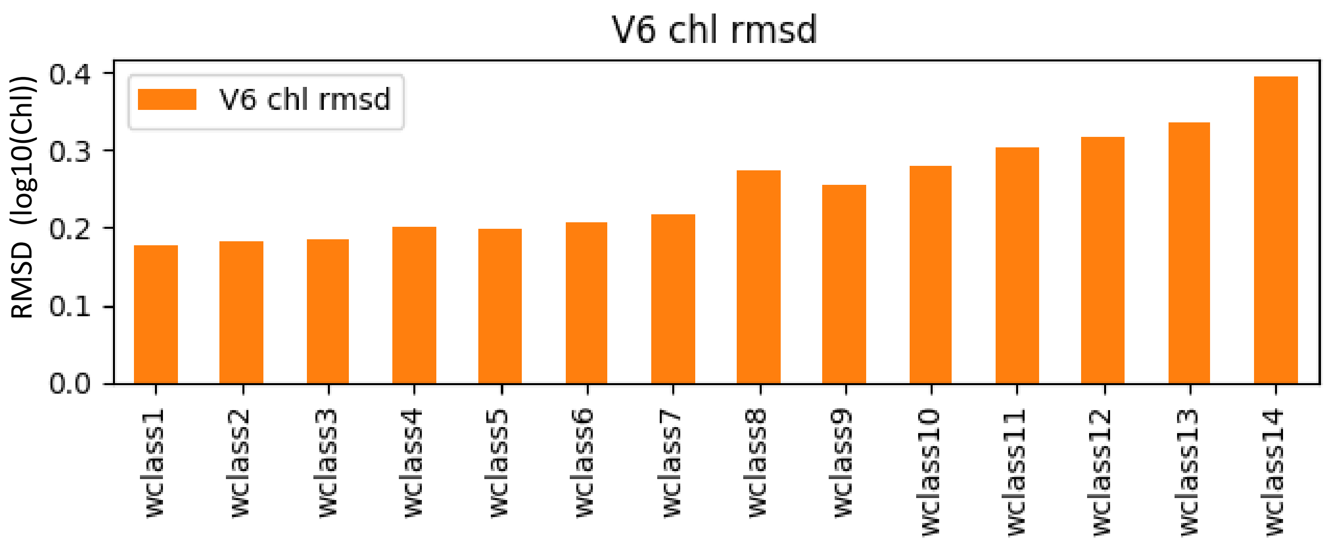

Figure 11: Chlorophyll-a average RMSD (log10(chl)) per optical water class (wclass).

Figure 11 shows that there is an almost unidirectional trend of increasing RMSD with increasing water class number. The clearest ocean waters have a relatively limited optical variability in comparison to coastal waters that are more optically-complex. It is therefore not surprising to see this trend in RMSD across the water classes as the more optical complexity the harder it is to separate the contribution of a single component, such as chlorophyll-a, from the total optical signal detected.

For the Rrs products we also have to consider the spectral nature of the data. The plots shown in figures 12 and 13 show the same statistics as for chlorophyll-a (bias and RMSD per water class) but also separate the results by wavelength.

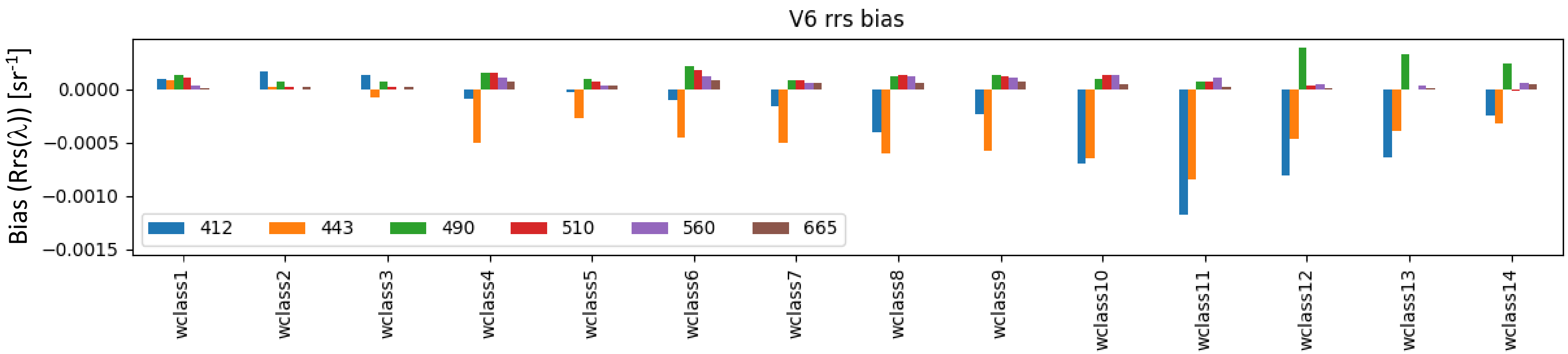

Figure 12: Rrs bias per optical water class (wclass).

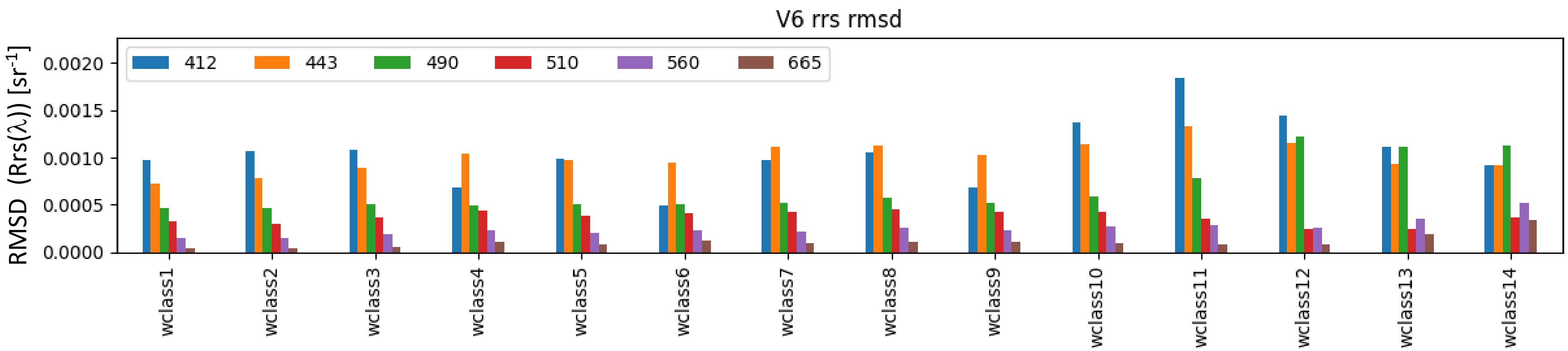

Figure 13: Rrs RMSE, labelled as root- mean-square difference (RMSD), per optical water class (wclass).

In the lower water classes (oligotrophic waters) there is a small bias across all wavelengths. As we move into the more turbid and coastal water classes (higher water class numbers) the bias becomes negative and the largest biases are at 412 and 443 nm. It is notable that the Rrs at 510 nm seems to possess a low bias across almost all the water classes.

Given that the Rrs data was used as input for the chlorophyll-a algorithms, it might be expected that the observations noted for the chlorophyll-a bias would be merely a reflection of the Rrs uncertainties. However, the chlorophyll-a biases are smallest in classes 3, 4, 5 and 13 rather than 1, 2 and 3 as we see in the Rrs. This is probably due to the fact that the strength of the signal from chlorophyll-a is so low in the clearest waters that even a small bias can have a relatively high impact.

The spectral distribution of RMSE for Rrs in most waters is high for shorter wavelengths and minimal for longer wavelengths; the most turbid waters possess more uniform RMSE values across the spectrum. These sorts of per water class differences in performance highlight areas that might be suited for targeted improvement in algorithm performance.

2.3.4 Comparison to the previous version

Both the overall and per water class results show an increased product performance when compared to the V5.0 dataset. The V6.0 Chlorophyll-a data has an increased correlation coefficient, a slightly reduced root-mean-square difference, an improved slope of regression and an improved intercept.

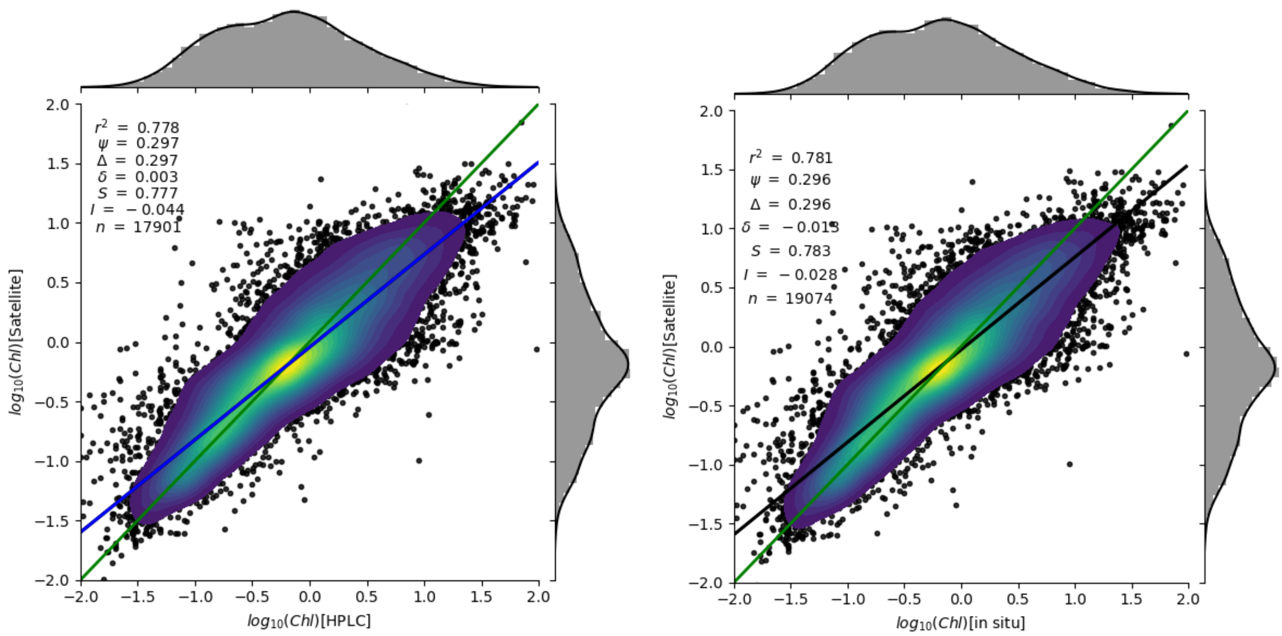

Figure 14: Comparison of the V5.0(left) and V6.0 (right) chlorophyll-a overall performance plots. The same display criteria used in figure 9 are used here, with the exception that in the V5.0 plot, the linear regression is shown in blue, with the V6.0 shown in black.

It is also of note that the new timeseries is not only extended, but now includes data from the OLCI sensor aboard Sentinel 3A and 3B. This has dramatically increased the coverage in recent years (as can be seen in **Figure 16**). The introduction of Sentinel 3B, which was not used in V5, is seen in 2018. However, the coverage then falls after 2019 as MODIS and VIIRS are removed from the V6 processing chain. Despite the removal of VIIRS and MODIS from 2020 onwards, the addition of Sentinel 3B has meant an overall increase in coverage in recent years.

Though the OC-CCI in-situ database is a global compilation, it contains a number of long-term time series measurements. One such time series is the Marine Optical BuoY (MOBY) funded by NOAA and located in the clear ocean waters near Hawaii. This is a very well maintained and high quality dataset which includes reflectance values for use in satellite validation. We extracted a regional timeseries around the MOBY station and compared it to the in-situ timeseries (see **Figure 15**). Looking at the V5 and V6 timeseries we can see that both agree well with the in-situ measurements. The V6 data seems to show a little higher variance than the V5, particularly in the earlier part of the record. However, the range of values are in good agreement with the in-situ data and the V6 data certainly seems to contain lower Rrs560 values (which are present in the in-situ data) than the V5 data. The V6 record appears to be stable (no drift away from the in-situ timeseries).

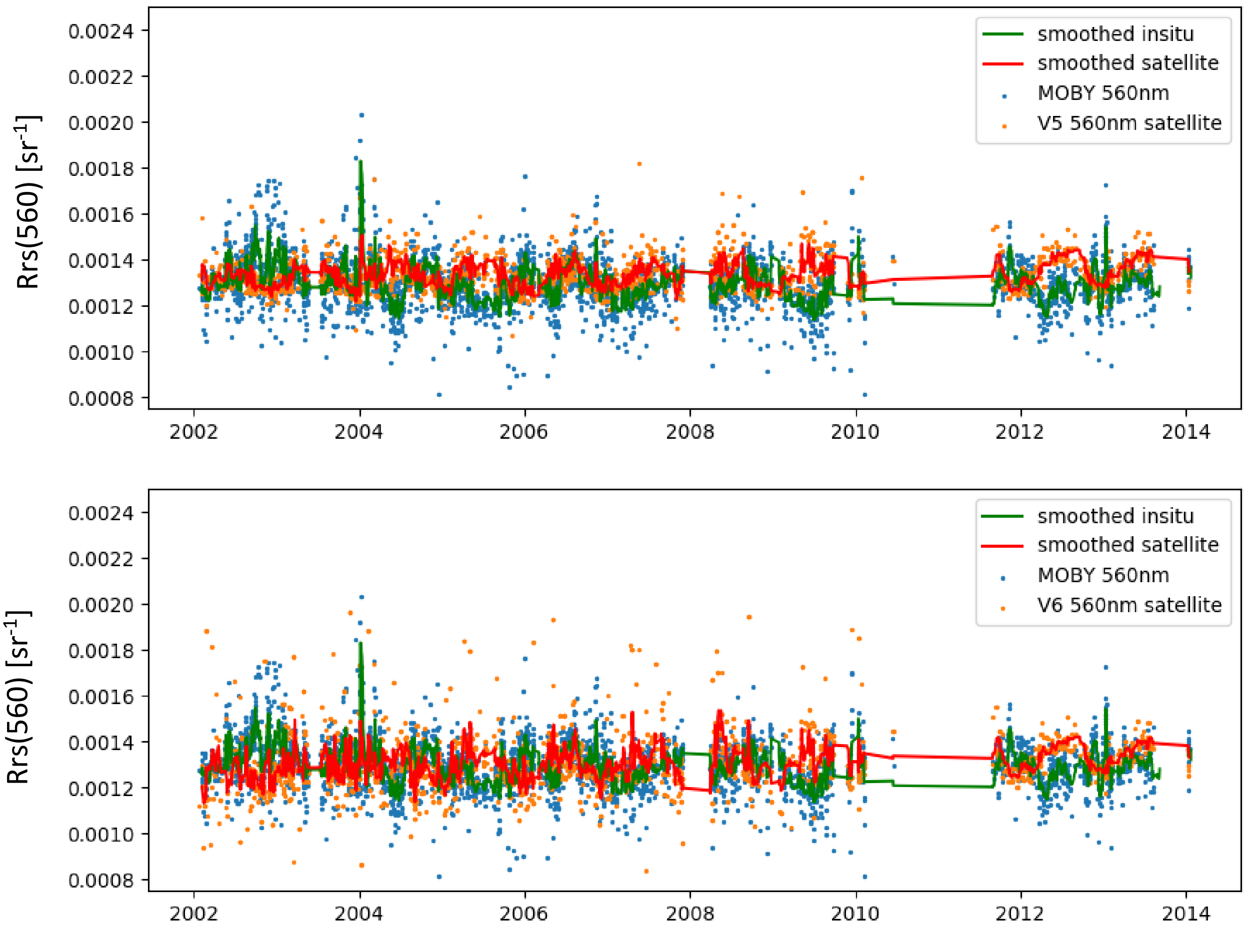

Figure 15: Comparison of a regional satellite time series Rrs (centered on MOBY location) and the in-situ timeseries from the MOBY station. 560nm band information is shown, and the rolling window used for data smoothing is 10 days, where at least 3 must provide data.

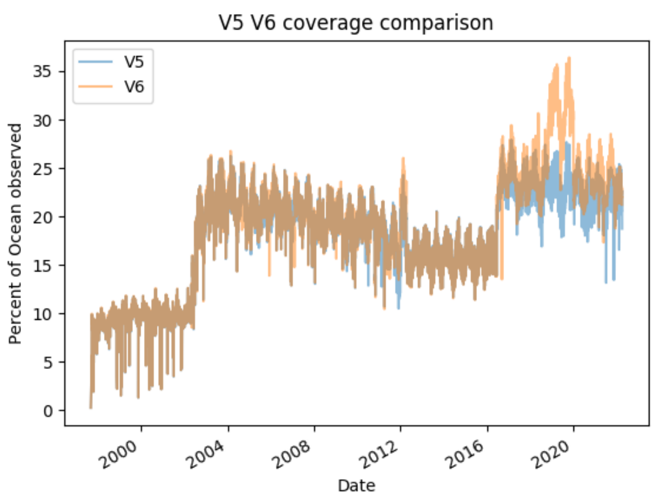

Figure 16: Daily percentage of the global oceans observed through time for the V5 and V6 products. Note that transparency is used due to large area of overlap between the 2 records.

3. Application(s) specific assessments

No application(s) specific assessments have been conducted as part of the work underpinning this PQAR. However, the authors would encourage users of the data to contact the C3S with feedback on their use of the data, for consideration in future advisories on application suitability.

The C3S Copernicus User Support (CUS)[10] is the point of contact for feedback on using the data.

[10] https://climate.copernicus.eu/help-and-support [last accessed 23rd January 2023]

4. Compliance with user requirements

User requirements for the OC-CCI dataset were determined within the scope of the OC-CCI project. For details of compliance of the dataset with these user requirements, see the [OC_CCI-CECR] document.

It is clear that in terms of temporal and spatial resolution the ocean colour products provided in the C3S catalogue are able to meet the community requirements (as summarised in Tables 1 and 2 by the GCOS requirements).

Table 1: Water leaving radiance requirements table comparing GOCS requirements to provided datasets.

GCOS requirement		**Currently achievable performance (as delivered in V6.0 products)**
Accuracy	5 % Specifically for the blue and green wavelengths in Case-1 Waters	5-15 % for water leaving radiances (for the blue and green wavelengths)
Stability	0.5 % per year	Has not yet been assessed
Spatial resolution	4 km (global) 1 km or smaller for regional and coastal applications	4-9 km horizontal 1 km horizontal resolution is available
Temporal resolution	Daily	Daily, weekly, monthly observing cycles are available globally, dependent on sensor

Table 2: Chlorophyll-a concentration requirements table comparing GOCS requirements to provided datasets.

GCOS requirement		**Currently achievable performance (as delivered in V6.0 products)**
Accuracy	30 % in Case-1 waters in the concentration range 0.01-10 [mg m-3]	5-50 % in Case-1 waters. 60-70 % for coastal waters and regional seas, which are typically Case-2. In areas of extreme optical-complexity these errors can be as high as 200-300 %. For these areas it is recommended that tailored algorithms be implemented.
Stability	3 % per year	Has not yet been assessed.
Spatial resolution	30 km (global)	4-9 km horizontal 1km is possible

Temporal resolution	Weekly averages	Daily, weekly, monthly observing cycles are available at global scale, dependent on sensor.

The bias information shown above (when mapped through the water classes onto the final products) shows that the uncertainty for chlorophyll (estimated from relative errors) is within the GCOS 30% accuracy requirement for much of the global oceans. Meanwhile most of the Rrs products show a bias averaging around 5%, except for the green band. Nevertheless, further improvements are underway, including better characterisation of coastal waters; improved cloud and sea-ice characterisation; extending the time series with additional satellite data; and greater consistency in processing between sensors.

The comparison of product performance to GCOS requirements would be aided by an increased amount of in situ data in the clearest ocean waters (water classes 1-2) and multiple long time-series of high precision (for stability tests). Improvements in performance for optically complex waters is likely to result from algorithm development rather than increased in situ data (though new algorithms still require validation). As new algorithms are tested, and the resulting product improvement is quantified, then these algorithms can be inserted into the processing chain as has been done per ECV processing cycle thus far.

5. References

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. (2016) 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.

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: https://support.ecmwf.int/