ESA Ocean Colour Climate Change Initiative+ Phase 2

D1.1 User Requirement Document v3

Ref: D1.1 v3
Date: 15/12/2025

Document Properties

Project	:	**Ocean Colour Climate Change Initiative+ (OC_CCI+) – Phase 2**
Document Title:		**D1.1 User Requirement Document v3**
Reference	:	**D1.1 v3**
Date	:	**15 December 2025**
Issue	:	**Final**
	:	**ESA / ESRIN**
Lead Author	:	[Dr Samantha Lavender – Pixalytics Ltd]
Science Lead	:	[Prof Steve Groom – PML]
Reviewed	:	[Prof Steve Groom – PML – PML]
ESA Technical Officer		[Dr Sarah Connors – ESA] Digitally signed by Sarah Connors Date: 2026.01.05 09:04:25 Z

Address : Plymouth Marine Laboratory (PML)
Prospect Place
The Hoe
Plymouth, PL1 2DH
United Kingdom

Tel: +44 (0)1752 633100

Table of Figures

Table of Tables

1. Summary

- The Ocean-Colour Climate Change Initiative (OC-CCI) project has been developing and releasing a merged time series of ocean colour products derived from multiple satellite missions since 2010 aiming to meet the key user requirements, expressed through the Global Climate Observation System (GCOS), of long-term, reliable, quality-controlled, stable products. The latest release is version 6.0 and work is underway in the CCI Extension (CCI+) Phase 2 –termed CCI+ phase 2, to further develop the products. To guide product development and delivery to the community, user requirements have been gathered and reviewed regulalrly throughout the series of OC-CCI projects.
- The 2015 Phase 1 user survey confirmed that stability was critical for trend analysis in climate studies. The need to establish improved stability requirements for time-series data analysis was also emphasised in the Phase 1 IOCCG Review, which pointed to the need for further work along the lines of Henson et al. (2010). Furthermore, applications of seasonal variability (phenology) tended to be insensitive to accuracy and depend on relative changes across seasonal and annual cycles. For these, high-frequency observations (weekly or more frequently) were considered essential. At the same time, identifying ecological provinces required high precision to enable statistical analysis. There was also a general interest in additional products that work seamlessly across Case 1 and Case 2 waters – now termed Oceanic and Optically Complex Water types. Therefore, a key goal of Phase 2 was to improve the uncertainty in Optically Complex Waters. During Phase 3, the helpdesk received user support requests for guidance on downloading subsets, receiving more frequent updates, and accessing companion datasets that supported improvements in data access.
- In CCI+ phase 2 the user survey has been conducted alongside the new ESA PHYTO-CCI project to extend the gathering of information to support the ongoing R&D activities. The dataset remains a key input to the community, supporting many publications each year (see https://climate.esa.int/en/projects/ocean-colour/), and there continues to be a need to improve and enhance the contents, documentation and availability inline with user expectations.

2. Introduction

The oceans play a critical role in the global carbon cycle, taking up approximately 25% of the carbon dioxide emitted by anthropogenic activities since 1850 (Hauck et al., 2020). Before the start of the first OC-CCI project, it was recognised that anthropogenic climate change was altering marine ecosystems at an unprecedented rate, but also in ways that push these systems outside their natural range of variation (IPCC, 2014). Therefore, it was concluded that the role of the oceans and the effects climate change may have on them and the carbon cycle must be understood.

Various measurement and modelling approaches can be used to evaluate the carbon pools in the oceans, with the carbon pumps recognised as being more complex than initially considered (Claustre et al., 2021). In situ sampling, satellites and numerical models, each provide different information to aid our understanding of the dynamics and variability of carbon pools and fluxes.

While in situ data are often used to provide ground-truth datasets for the calibration and validation of satellite data and parametrization of numerical models, in situ data collection is labour intensive and can be expensive. Satellite Ocean Colour (OC) data can provide global datasets, but the derived variables are limited to those detectable or inferable from their optical signature. Furthermore, OC data can be restricted in time and space because of poor retrievals at high solar zenith angles, interorbital gaps and the presence of aerosols and clouds that obscure the ocean surface. Finally, Earth System models represent the best approximation of processes driving the carbon cycle and allow for a global representation where processes are sufficiently understood. However, such models rely on in situ and remote sensing observations to parameterise the impact of ocean warming, ocean acidification, and deoxygenation on marine ecosystems alongside global biogeochemical cycling (Gruber, 2011).

Satellite OC assimilation can improve numerical models. In turn, numerical models can support the planning and design of field campaigns by simulations and forecasts of what could be encountered during the sampling campaign. Climate and marine ecosystem model inter-comparison projects, such as the Coupled Model Intercomparison Project 5 (CMIP5) and MARine Ecosystem Modelling Intercomparison Project (MAREMIP) (Taylor et al. 2012; Vogt et al. 2013), have shed light on gaps in our understanding of future ocean conditions and our ability to quantify them, including the role of potential climate feedbacks. While the physical changes due to global warming are consistently projected by models, with a medium-to-good understanding of surface ocean temperature and pH changes, the uncertainty in ocean productivity and carbon export has often been substantial (Bopp et al., 2013).

Ocean Colour was identified as an ECV by the Global Climate Observation System (GCOS), within the GCOS Implementation Plan (GCOS, 2004), identifying an initial set of ocean-colour requirements for climate research (see Appendix 1 of the implementation plan for a summary of GCOS requirements, and Appendix 2 for an overview of the status at that time). The systematic observation requirements were subsequently identified (GCOS, 2011) and, then new Implementation Plans were released GCOS (2016 and 2022). In 2024 GCOS started a process aimed at the rationalization of the ECV list. Based on feedback, the final list will come into force when the new GCOS Implementation Plan is published (2028).

The Ocean Colour Climate Change Initiative (OC-CCI) aims to improve ocean-colour products through the merging of ‘climate quality’ missions. OC is one of many Essential Climate Variables (ECVs) accessible to remote sensing that have been studied under the umbrella of the European Space Agency (ESA) Climate Change Initiative (CCI). Over the initial Phase 1 three-year period, 2010-2013, the OC-CCI scientists and collaborators developed, validated, and selected algorithms to generate the OC-CCI version 1.0 (v1.0) dataset. Since then, data releases in Phase 2 (2014 – 2017, extended to 2019) and Phase 3 (started in 2019) produced versions 2.0 through 6.0. The overall methodology, and versions up to 4.0, are described in Sathyendranath et al. (2019).

Chlorophyll-a concentration (Chl-a) is an indicator of photosynthetic biomass in significant regions of the ocean, so it was identified as an ECV; it is the single most crucial oceanic component responsible for changes in the signal detected by satellites (Gordon and Morel, 1983). Water-leaving radiance is the upwelling radiance just below the sea surface that is transformed into remote-sensing reflectance (Rrs), i.e. the ratio of the upwelling to downwelling light. Normalised water-leaving radiance (nLw) approximates the radiance that would leave the sea surface in the absence of an atmosphere when the sun is at the zenith.

2.1. Assessment Methodology

The OC user community's needs were assessed across several ESA-funded projects to define the required characteristics of satellite-derived biogeochemical products. In 2010, OC-CCI collected views from the scientific community using an online survey and workshop: the Phase 1 user survey. Then, a survey was conducted to understand the requirements of the modelling community and those scientists who use ocean-colour data directly to monitor decadal-scale variability in the marine ecosystem. The surveys aimed to identify current and near-future requirements for ocean-colour data for climate studies; the user requirements specified by GCOS (2004) and the Climate Model User Group (CMUG) of ESA served as the starting points. The goal was to improve interactions between the ocean-colour community and users, particularly the modelling community. Therefore, the survey included questions to evaluate how well ocean-colour products meet climate research needs and establish priorities for further development to enhance user engagement. These inputs fed into Phase 1 and the production of the OC-CCI v1.0 dataset.

Specifically, respondents were asked about product spatial and temporal resolution, data access, stability, and uncertainty. Requirements were divided into "threshold" (minimum requirement to be met to ensure that data are useful), "goal" (an ideal requirement, above which further improvements are not necessary) and "breakthrough" (an intermediate level between "threshold" and "goal" which, if achieved, would result in a significant improvement for the targeted application) with uncertainty separated into accuracy and precision:

- Accuracy: a measure of the non-random, systematic error or bias that defines the offset between the measured value and the true value that constitutes the SI absolute standard.
- Precision: a measure of reproducibility or repeatability of the measurement without reference to an international standard, so that precision is a measure of the random and not the systematic error. Suitable averaging of the random error can improve the precision of the measurement, but does not establish the systematic error of the observation.

During Phase 2, the initial approach was to cross-check and validate the user requirements that resulted in v1.0 of the dataset against the received feedback, both internally and externally. This approach included recording feedback during internal progress meetings, extracting feedback from CMUG review documents (e.g., CMUG, 2010), analysing direct user enquiries via the helpdesk email address, and an informal workshop with the Climate Research Group (CRG).

Systematic observation requirements had also been identified by the GCOS (2011), and so a Phase 2 (2015) survey built upon the Phase 1 activity. This survey was a joint activity of the ESA OC-CCI, Pools of Carbon in the Ocean (POCO) and Photosynthetically Active Radiation (PAR) projects. Respondents were again separated into modellers and Earth Observation (EO) scientists.

During Phase 3, the focus of the time-series has been on including further missions and reducing the error sources. The version 5.0 dataset merged data from MERIS, MODIS-Aqua, SeaWiFS, Sentinel-3A OLCI and Suomi-VIIRS. A significant change from previous versions, was that version 5.0 updated the reference sensor from SeaWiFS to MERIS with a future move to OLCI planned.

The latest version, 6.0, was released in November 2022. It included Sentinel 3B OLCI data and the MERIS 4th reprocessing data . Through a combination of ESA and UK national funding version 6 was extended with three month Intermediate CDR releases up to the end of March 2025 when the current CCI+ phase 2 started.

During this CCI+ Extension, the survey at the start of this phase has been conducted as a collaborative assessment with workshops to gain more in-depth insights. This deliverable will be updated as part of the ongoing user engagement activities and in support of the next version.

2.2. Summary of OC-CCI Data Usage

Many EO scientists are using OC data directly in climate studies. For those using trend analysis, stability continues to be critical. Ohring et al. (2005) suggested that stability should be at least 20 % of expected trends and, for OC, should be 1 % over a decade and reported that OC was close to meeting that requirement. The need to better establish the stability requirements for time series data analysis was also emphasised in the Phase 1 IOCCG Review. This review pointed to the need for more work along the lines of Henson et al. (2010), who used models to evaluate the length of data records necessary to detect the effect of climate change on OC. Since then, Van Oostende et al. (2022) studied the OC-CCI dataset and showed that coastal waters, high latitudes, and areas subject to changing cloud cover are most affected by coverage variability between missions.

Some applications of OC, such as studying variability in seasonal dynamics (phenology) of phytoplankton, are insensitive to accuracy in retrieved products but depend on relative changes over seasonal and annual cycles. A high frequency of observations (weekly or better) was deemed essential for such applications to enable timings of rapid changes in the ecosystem, such as the onset of the spring bloom. Based on a modelling study, Henson et al. (2018) found that in response to projected 21st century climate change, bloom timing generally shifted later at mid-latitudes and earlier at high and low latitudes by ~5 days per decade to 2100.

Another application of OC in climate studies is the identification of ecological provinces in the ocean: remote sensing is the most effective method for delineating these provinces' boundaries and tracking changes in the boundaries with time. Different methods are used to identify provinces and their boundaries by remote sensing, and requirements vary with the technique. Still, high precision would facilitate statistical techniques such as clustering to distinguish one province from another.

It was desirable to have algorithms that work seamlessly across Oceanic and Optically Complex Water types as a long-term goal. Therefore, a key goal of Phase 2 was to improve the uncertainty in Case 2 waters. The change in approach included using the optical classification (Jackson et al., 2017) to distinguish between different water types and apply different Chl-a algorithms (Jackson, 2020). This has improved the products, but there remains further work to ensure consistency as previously noted in Van Oostende et al. 2022).

During Phase 1, there was also a general interest in new products from OC, such as information on particle size structure, phytoplankton functional types, and measures of concentrations of coloured dissolved organic matter. As overviewed within the Colour and Light in the Ocean from Earth Observation (CLEO) workshop (Sathyendranath et al. 2016), detailed in Section 3.4.4, additional products were being addressed by projects outside of OC-CCI but needed to be closely linked. This is continuing with projects such as ESA PHYTO-CCI.

A summary of the usage by the climate community is provided within the Climate Assessment Report, last released in 2024 (see https://climate.esa.int/en/projects/ocean-colour/) and this will be updated in the CCI+ extension.

3. Results

3.3. User Surveys

3.3.1. Phase 1: 2011

The Phase 1 user survey was conducted via a web-based multiple-choice questionnaire in which respondents were asked to provide brief justifications. Responses received were sorted by country and the two main user groups were targeted: climate modellers (shortened to modellers) and EO scientists; see **Figure 3-1**. As most respondents were from the EO community, they were mainly interested in global-scale analysis, primary production studies, ecological provinces and fisheries variability, with additional interests including photochemistry, bio-optical studies, Carbon cycle studies and water quality monitoring. Climate modellers were interested mainly in OC for model validation/skill assessment. There was again a split in spatial-scale requirements, depending on regional or global applications.

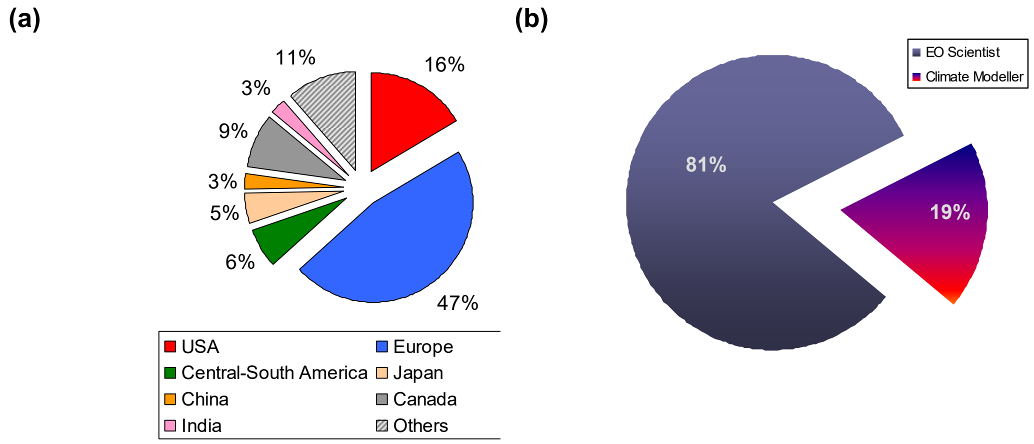


Figure 3-1: Partitioning of the Phase 1 user survey by (a) country and (b) user groups targeted.

The main product from OC is Chl-a and given the development of techniques to derive other biogeochemically-significant variables from the ocean-colour signal, users were also asked about their interest in these other remote sensing products. Some products were routinely obtained, such as nLw and the spectral attenuation coefficient for downwelling irradiance (K_d). In contrast, others had been developed more recently or were experimental but will likely be validated products in a horizon of 5 years. These variables were not explicitly defined in the GCOS requirements, and, therefore, there were no specific requirements for them at the time.

Table 3-1 summarises the questionnaire results related to uncertainty and stability, with the total number of responding scientists to the questionnaire being 78, 15 climate modellers and 63 EO scientists.

The "threshold" requirement levels for the accuracy, precision and stability are reported for Chl-a alongside nLw, K_d, inherent optical properties (IOP), coloured dissolved organic material (CDOM), suspended particulate matter (SPM), phytoplankton functional types (PFT), and particle size distribution slope (PSD). Stability was also included as it is vital for long-term records. The bias defines the time-dependent (or instrument-dependent) difference between the observed quantity and truth when no absolute standard is available to establish the systematic error quantitatively. The questionnaire also highlighted a clear requirement that uncertainty estimates for each product should be included as a standard.

Breakthrough accuracy requirement levels from the CMUG were similar to those that emerged from the questionnaire and were more stringent than the original GCOS (2011) breakthrough requirements but in line with the revised GCOS (2016) requirements. Modellers required an accuracy of 10 – 25 % for the precision of Chl-a. A request is based on the needs of unbiased validation data sets and data sets for assimilation, as well as an awareness of the capabilities of current satellites; this was at variance with the view of many of the experts who attended the London meeting, who considered it was more important to have high accuracy than high precision (Martinez Vicente, pers. comm.). For EO scientists, the accuracy requirement is 10 – 25 % as it avoids that the bias propagates into derived variables such as primary production or gas-flux computations. Precision requirements had not been explicitly recorded in the GCOS or CMUG reports. Still, the expectation of modellers and EO scientists is likely to be a similar precision to in-situ measurements of Chl-a (~10%).

The stability of the monitoring of the ECV is a characteristic that is crucial for climate change studies. There was a similar request from modellers and EO scientists. For modellers, 10 % variation over ten years would be considered sufficient for data assimilation studies, whilst a 5 % variation over ten years would be a maximum to allow for climate studies. EO scientists believed that given recent publications at the time (e.g.

Boyce et al. 2010), where an annual 1 % decline in Chl-a was observed, stability of at least 1 % over ten years would be required for detecting climate change.

The users were also asked about the importance of having error estimates assigned to Chl-a derived from EO. Both communities had a very clear requirement, requesting product error estimates to be made available along with the product values (100 % of modellers and 95 % of EO scientists believed having error estimates was essential).

Table 3-1 Summary of the uncertainty and stability "threshold" requirements from the Phase 1 survey alongside values taken from GCOS (2011).

Parameter	User type	Accuracy (%)	Precision (%)	Stability (% change $10y^{-1}$)
Chlorophyll	GCOS (2011)	30		3
	Modellers	10-25	10-25	10
	EO Scientists	10-25	10-25†	5
nLw	GCOS (2011)	5**		0.5
	Modellers	10-25	Around 25	2
	EO Scientists	Better than 10	Better than 10	1
K_d	Modellers	10-25	10-25	2
	EO Scientists	Better than 10	Better than 10	1
IOP	Modellers	10-25	10-25	Around 2
	EO Scientists	Better than 10	Better than 10	Around 2
CDOM	Modellers	25-50	10-25	10
	EO Scientists	Better than 10	Better than 10	5
SPM	Modellers	10-25	10-25	5
	EO Scientists	10-25	Better than 10	2
PFT	Modellers	10-25	10-25	2
	EO Scientists	10-25	10-25	5
PSD	Modellers	10-25	10-25	2
	EO Scientists	10-25	Around 25	5

*** The 5% requirement was specifically for the blue and green wavelengths.*

Concerning the horizontal resolution required for Chl-a, the requirements from the EO scientists were more stringent than those from GCOS or CMUG. CCI modellers' requirements (Goal and Breakthrough) agreed with CMUG estimates. The important feature for this group of users was to obtain satellite Chl-a data at a resolution similar to global ocean models (grid size of ~10 km or coarser). The reasons expressed in the questionnaire for the EO community to request 0.1 – 1 km resolution were to capture small-scale processes in coastal studies of climate change and to incorporate sub-mesoscale processes into global models.

The observing cycles (temporal frequency of sampling) required was one week compared to one day from CMUG and one week from GCOS. Weekly Chl-a data at a resolution of at least one week was required to account for the limitations imposed by cloud cover. The monthly resolution would be sufficient for model evaluation, but the weekly frequency (at least) was deemed critical for data assimilation. However, any product averaged over longer periods than one week would miss the timing of phytoplankton blooms completely and hence would not allow the study of seasonality and changes thereof.

Figure 3-2 shows the distribution of interest for the two groups of users across the different biogeochemically-significant variables. Both groups show similar interest in all proposed variables, with more interest in nLw by EO scientists who were better equipped to handle this radiometric product. For this question, there were significant requests for providing primary production and PAR with other requests for inorganic or calcite concentrations and particulate organic carbon within the free field. Additional parameters comparable with historical optical measurements, such as Secchi disk depths and the Forel-Ule scale, were requested, alongside zooplankton estimates.

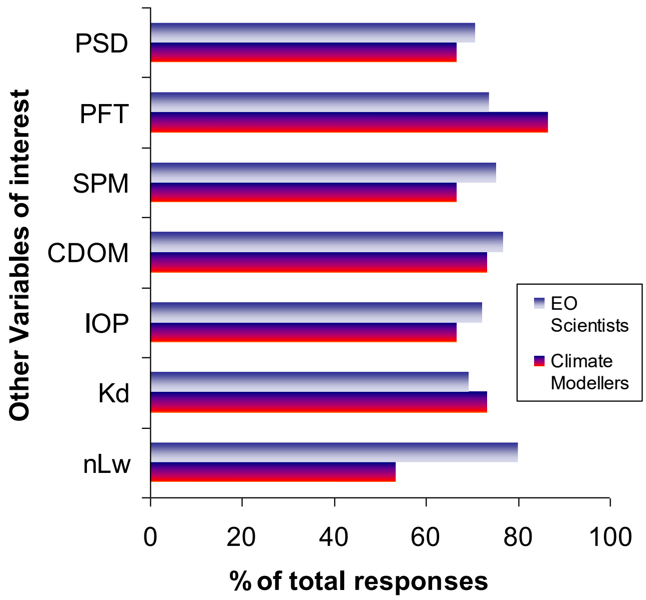


Figure 3-2: Percentage of other variables of interest for EO scientists and Climate modellers.

The questionnaire was also dedicated to data format and access preferences. Users were questioned about the level of processing required for their data, and were given the following (multiple-choice) options:

- Level 1 Geophysical Measurements, i.e. radiances;
- Level 2 Datasets, i.e. derived products on original space view;
- Level 3 Datasets, i.e. daily and monthly mean gridded.

The results indicated significant differences among both users communities; see **Figure 3-3**. Most EO scientists preferred to obtain Level 1 data to test different algorithms (including atmospheric corrections and in-water algorithms). In contrast, modellers preffered to get Level 3 data extracted easily compared with models outputs or included in data assimilation schemes.

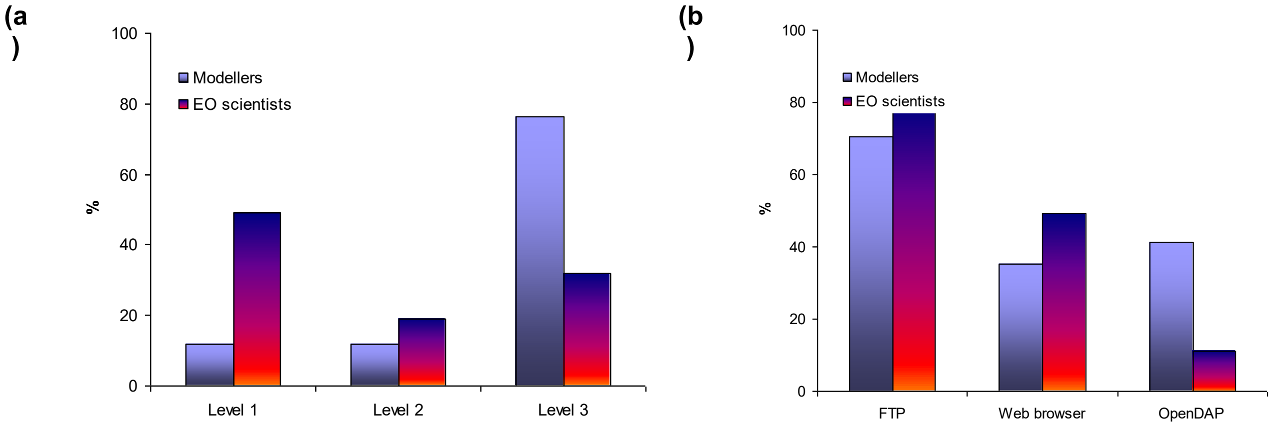


Figure 3-3: Preference of users concerning (a) level of data processing and (b) data access. As this was a multiple-choice question, hence the total number of answers to it was greater than the number of responses

Concerning the preferred means of data access, differences among users were less important. For format, there was a general agreement in preference of CF Compliant NetCDF, with some EO scientists also requesting data in HDF format, while modellers preferred data in the CIMP-5 format. Most of the users in both groups preferred FTP access, and then EO scientists preferred Web browser access while modellers preferred OPeNDAP.

In terms of tools for analysing remote sensing data, modellers tended to use Matlab and R as a primary tool. In contrast, EO scientists used, apart from Matlab and R, mainly used SeaDAS, BEAM, IDL, Python and ERDAS, as well as SISCAL and ARC-GIS tools. Overall, most users prefered using free tools developed under an open-source license, which came with source code. Very few modellers (23 %) were aware of the open-source ESA BEAM Toolbox (http://www.brockmann-consult.de/beam/), in contrast with the EO scientists that were aware and use (or intended to use) BEAM (97 %).

3.3.2. Phase 2: 2015

Again, in 2015, the survey was carried out through a web-based multi-choice questionnaire. Still, this time the survey was part of a broader survey that also included an analysis of requirements for the ESA POCO and PAR projects. There were responses from the global communityÇ **Figure 3-4**. However, a large percentage of the respondents decided not to provide their details. So their country of origin could not be assessed, with 138 respondents declaring themselves as 81 EO scientists and 39 modellers. For further analysis, this had been split such that people who only declared their interest as only EO ended up in the EO category, and EO plus modelling ended up classified as a modeller.

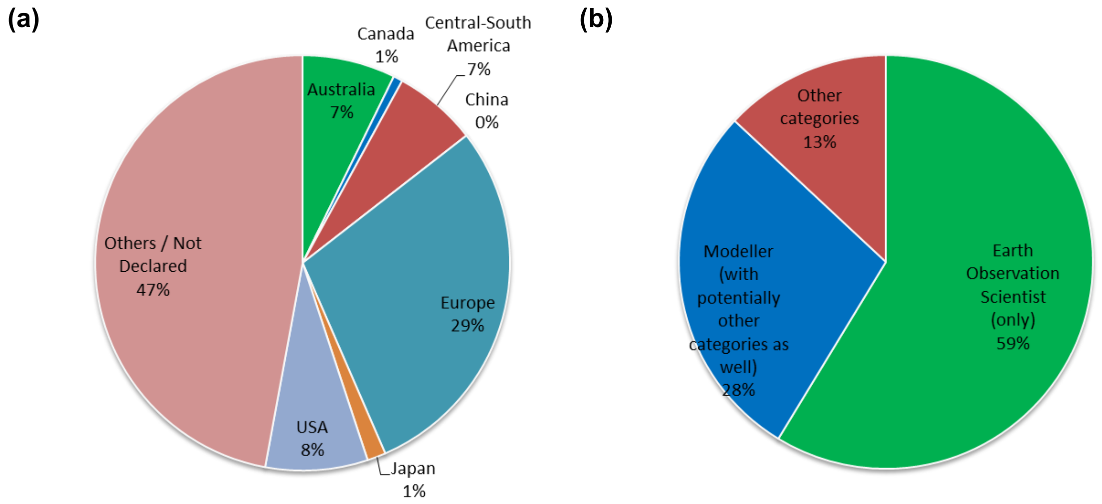
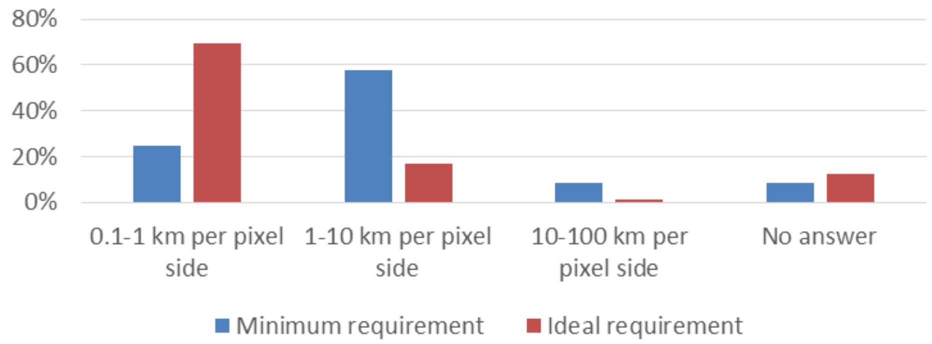


Figure 3-4: Partitioning of the Phase 1 user survey by (a) country and (b) user groups targeted.

The focus was on the contents of the released OC-CCI dataset, and so the questions asked related to the dataset products: Rrs, IOP, Chl-a and Kd. Of the users sampled, 94 % use satellite-derived ocean-colour data, and 93 % were using or interested in using the OC-CCI products; with 27 % using the data to support modelling activities. The actual usage of individual product types was relatively evenly distributed but slightly higher for Chl-a: 19 % K; 20 % IOPs; 25 % Rrs; 36 % Chl-a.

A significant proportion of the users (70 %) would have preferred data with a spatial resolution of between 0.1 and 1 km, a higher spatial resolution than the standard OC-CCI products. However, the resolution of the products (4 km) did meet the majority's minimum requirement of 1 to 10 km resolution; see **Figure 3-5**. For the temporal resolution, the preference was for one image per day. Still, a coarser resolution of one image per week was an acceptable minimum requirement for the majority of the participants (82 %).

These findings were comparable to the Phase 1 survey where, for Chl-a (**Table 3-1**), the important feature was to obtain satellite Chl-a data at a resolution similar to global ocean models (specified as a grid size of ~10 km or coarser). For the EO community, 0.1 – 1 km resolution allows for capturing small-scale processes in coastal studies of climate change and incorporating sub-mesoscale processes into global models. The requirement for temporal frequency sampling was one week, as this can account for the limitations imposed by cloud cover. However, having a higher temporal resolution allows transitory features to be captured; it becomes more critical for coastal applications.



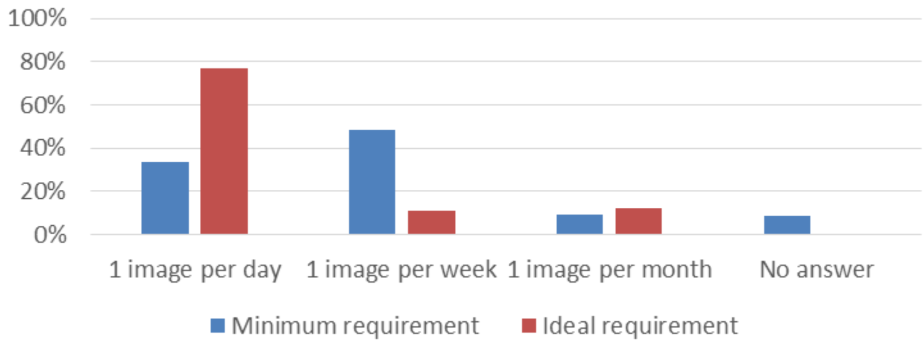


Figure 3-5: Minimum and ideal requirements in terms of (top) spatial resolution and (bottom) temporal resolution

All three OC-CCI data characteristics were considered essential, with an acceptable level of uncertainty being 25 % or less for all the OC-CCI products; see **Figure 3-6**. Also, most people responding to the question (75 % of people) felt having uncertainties was indispensable or desirable, while only 15 % classed them as not being required.

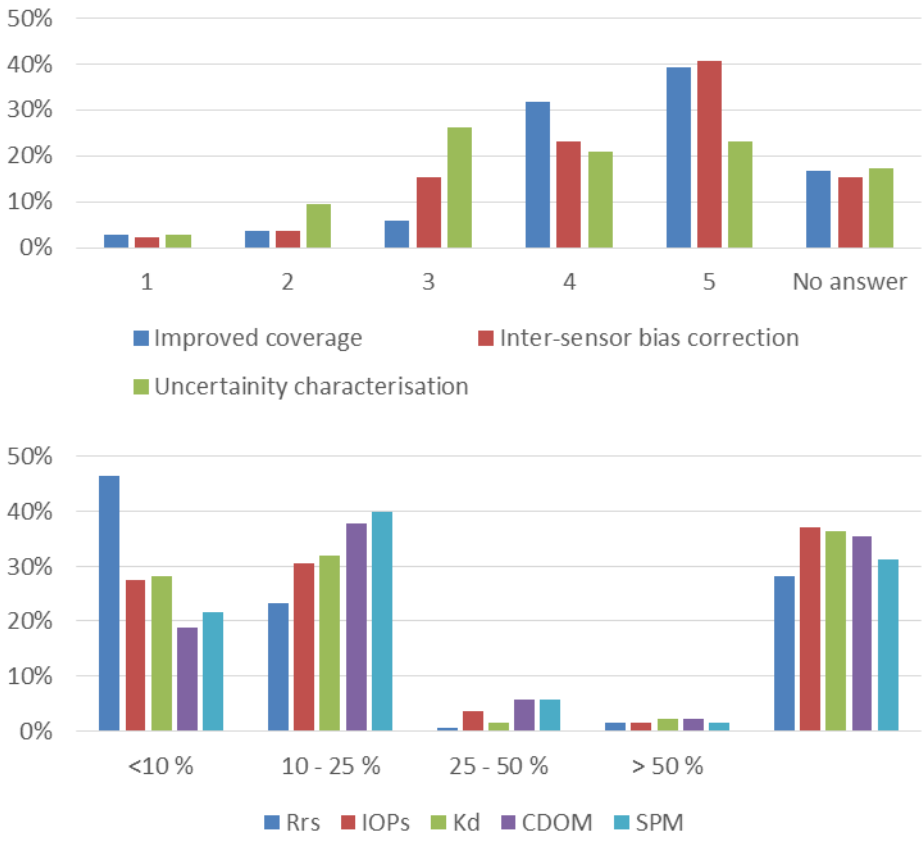


Figure 3-6: The (top) importance of the data characteristics and (bottom) the requirements in terms of an acceptable level of uncertainty.

Almost half (48 %) of the people responding to this question still downloaded data by FTP, with HTTP being the following preferred method (24 %) and then the Web GIS Portal (12 %). Approaches that allow on-the-fly sub-setting of the NetCDF files (OPeNDAP and THREDDS) accounted for 15 % together. An issue for the usage of THREDDS can be that downloads of greater than 4 GB as a NetCDF file were not possible via the oceancolour.org website; due to the default output format being NetCDF-3 rather than NetCDF-4. This limitation in download size became an increasing issue as the full v2.0 OC-CCI dataset volume was 46 TB, compared to 27 TB for v1.0. As an alternative, OPeNDAP allows a binary or ASCII output format of any size.

3.3.3. CCI+: 2025

The latest user survey was conducted in collaboration with the new ESA PHYTOplankton biomass and diversity Climate Change Initiative (PHYTO-CCI), which is developing phytoplankton biomass and diversity ECVs from the OC-CCI dataset. Therefore, the full survey findings are summarised in Kulk and Martinez-Vicente (2025), and this section summarises relevant elements for the OC-CCI project.

As with the previous user surveys, the aim was to understand the audience and to consider responses by the audience type that included both EO specialists and climate modellers; see **Figure 3-7**. Most users were EO scientists, and an additional "Other user" category included groups such as ecologists and water quality managers.

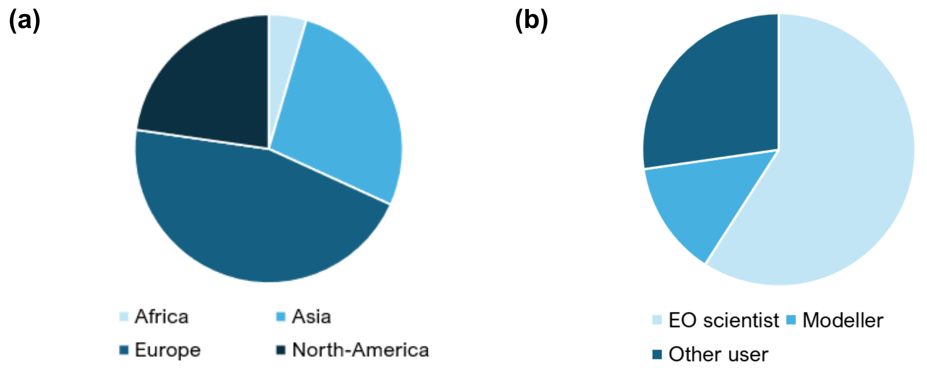


Figure 3-7: Partitioning of the CCI+ user survey by (a) region and (b) user groups targeted.

Users were asked about their perceived importance of having error characterisation, and the majority of EO users saw this data as important, while modellers were less sure; see **Figure 3-8**.

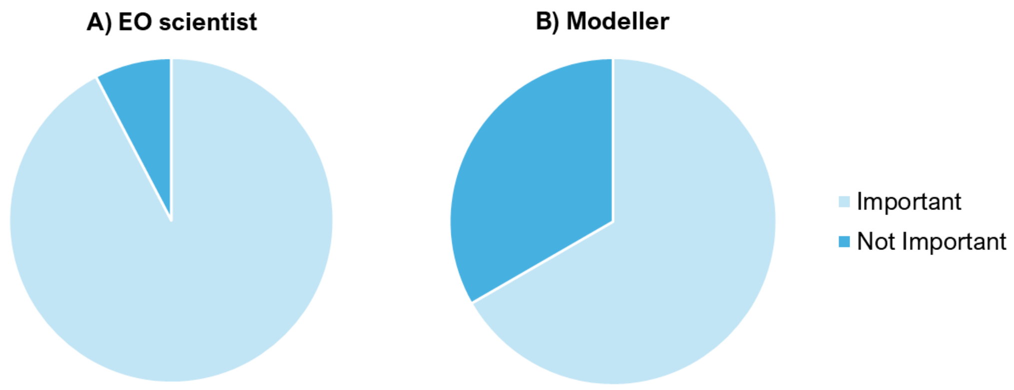


Figure 3-8: Importance of error characterisation.

The requirements for Chl-a regarding both resolution (horizontal and temporal) and error characterisation (accuracy, precision and stability) are shown in **Figure 3-9**, with separate bar charts for the different options when choosing the levels.

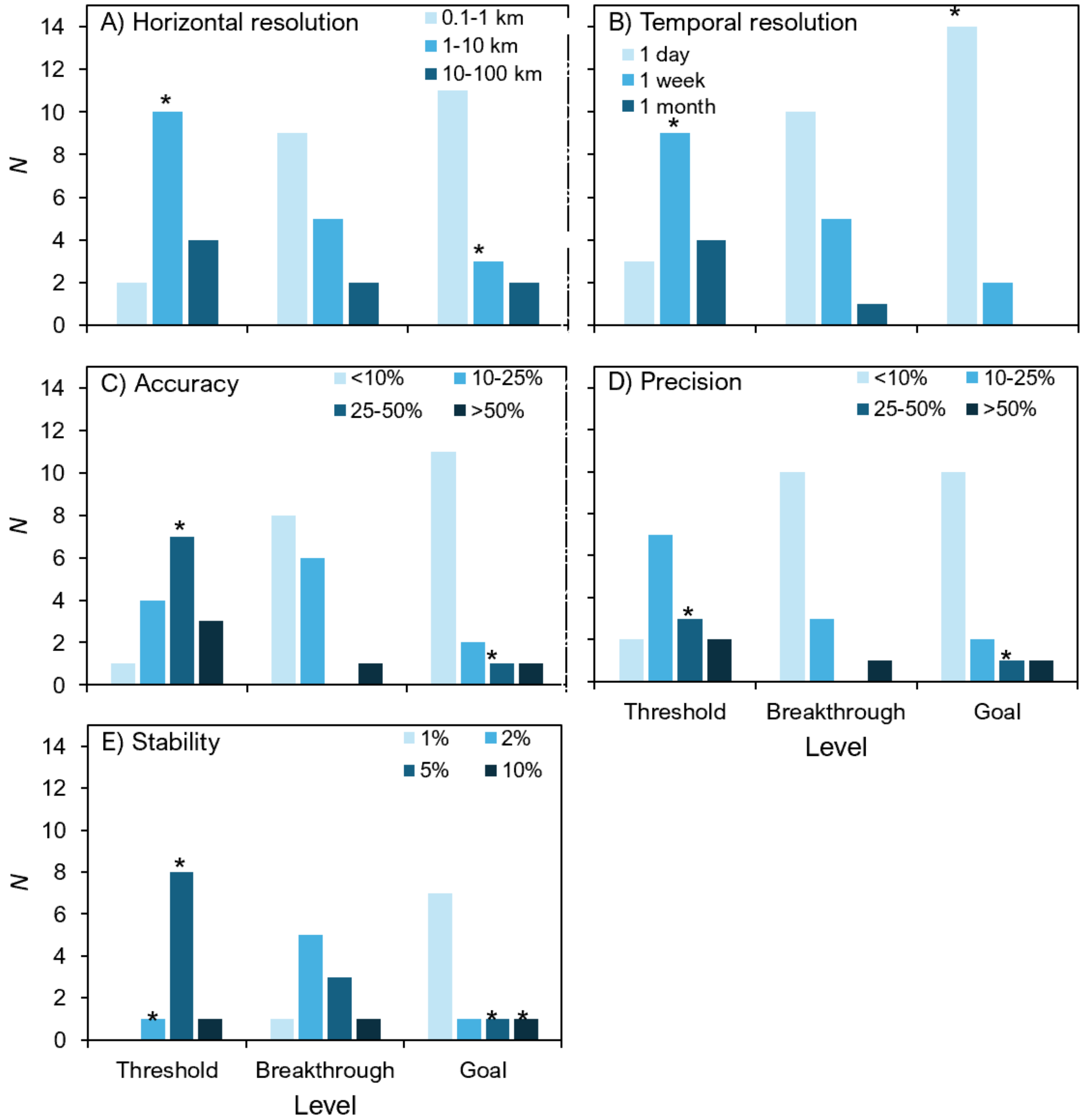


Figure 3-9: Chlorophyll-a requirements regarding (a) horizontal and (b) temporal resolution, (c) accuracy, (d) precision and (e) stability.

Data formats, access routes and access locations were surveyed to understand how and where from users were accessing the OC-CCI dataÇ see **Figure 3-10**. NetCDF continues to be the preferred format and there is still a strong wish for FTP alongside cloud-native methods, with the PML portal (oceancolour.org) and the Copernicus Climate Change Service (C3S) being trusted sources alongside the broader CCI website and CEDA where the long-term archive is held.

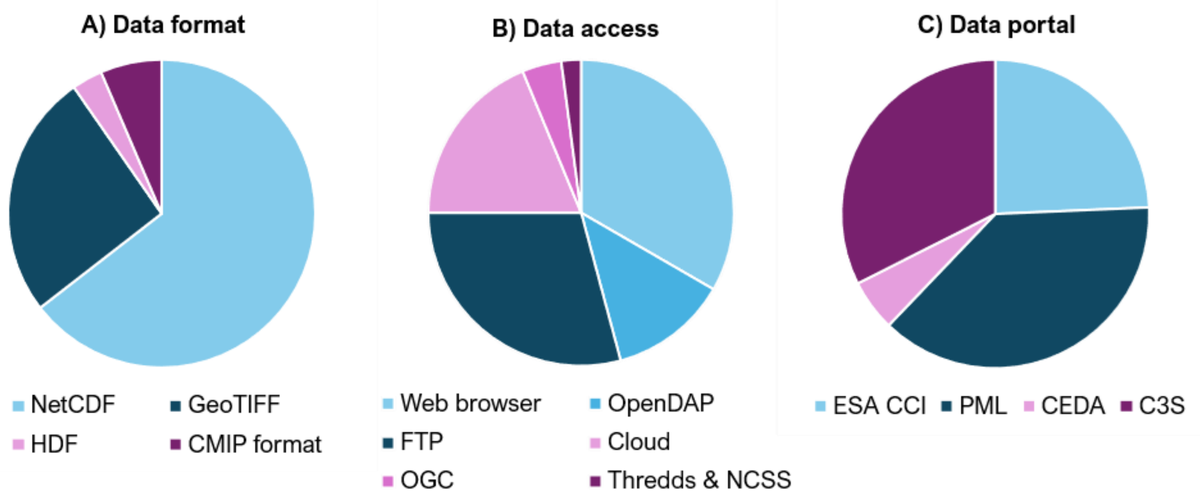


Figure 3-10: Summary of the preferred data formats, access routes and locations

As a comparison to the previous,Phase 1 survey, **Figure 3-11** shows the variables of interest beyond Chl-a. There is a lot of similarity, but also a strong interest in the derived products from those more focused on the applications of the data.

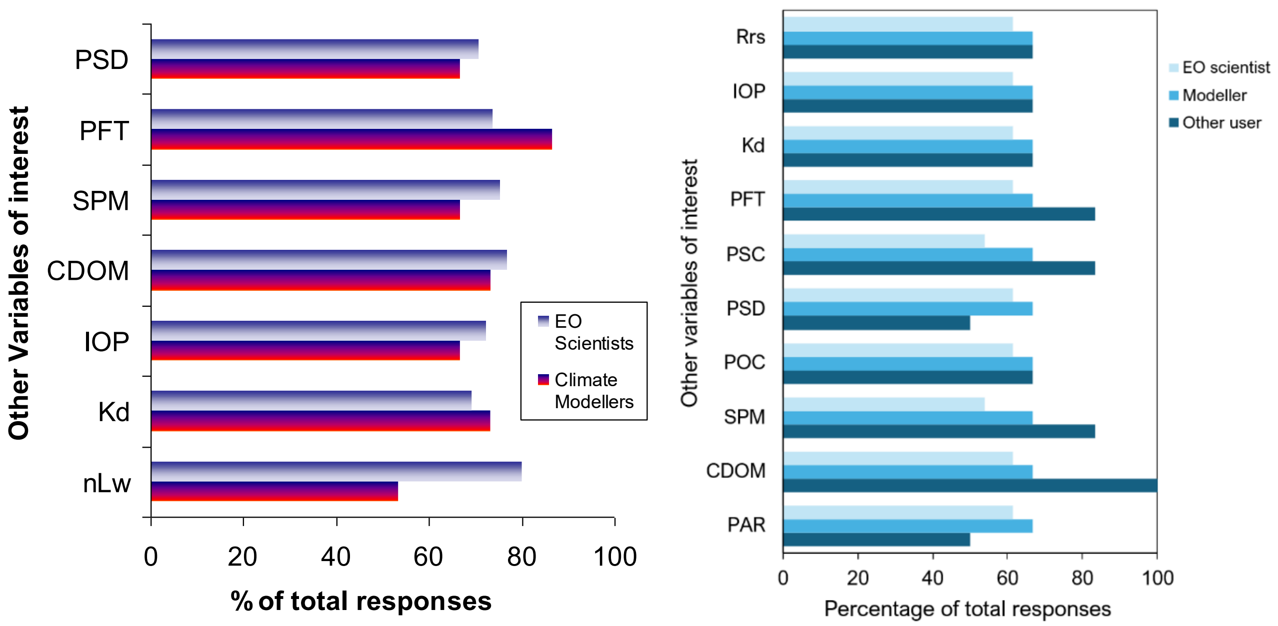


Figure 3-11: Variables of interest beyond Chl-a in 2011 (left) and 2025 (right).

3.4. Other approaches

3.4.4. Phase 2 & 3: Requirements gathering

The Colour and Light in the Ocean from Earth Observation (CLEO) Workshop was held in ESA (ESRIN) in September 2016. The objectives were to bring together users and providers of OC products from ESA Scientific Exploitation of Operational Missions (SEOM) projects, including OC-CCI. An outcome was the definition of a Scientific Roadmap (Sathyendranath et al. 2016) identifying challenge areas and research priorities for future EO data exploitation activities and shaping ideas for potential new OC products to be developed for climate research and ecosystem modelling in the era of the Copernicus Sentinel-3 mission.

The end-user requirements were summarised in the discussion session as being and including the curation of OC data because it was critical for the long-term detection of climate impacts. It was stated that datasets, such as those provided by OC-CCI, provide a time series that should be protected from gaps in the data stream; it was deemed essential to have at least two sensors in orbit at the same time. In addition, modifications were required in response to reprocessing, a system must accompany new sensors and advances in science plus these high-quality products to provide their access (user-friendly portals, user support, toolboxes, avoid misuse of data etc.).

In terms of validity and accuracy, the level of accuracy must be high so that users can detect minor trends. Also, merged OC products require a concurrent effort to maintain a selected set of long-term oceanic time series data for validation:

- Synergy with observations: bringing together in situ observations, satellite data and models to study ocean processes relevant to climate.
- Add in new autonomous and remotely-operated vehicles (gliders and bio-Argo)
- In situ validation and calibration programmes must keep pace with enhanced product lines.

The CMUG November 2021 Quality Assessment Report (CMUG, 2021) assessed the suitability of CCI ocean colour products to constrain biophysical feedback when assimilated into coupled physical-biogeochemical ocean reanalyses. Experiments performed (Ford, 2020) concluded that seasonal variations in the carbon-to-chlorophyll ratio play an essential role. Studies of phytoplankton bloom initiation based solely on chlorophyll data may not fully understand the underlying processes.

Also, during Phase 3, the helpdesk approach continued to be popular for receiving using feedback / enquires that included requests related to:

- **Downloading subsets of the dataset:** The THREDDS/OPeNDAP servers were popular, but restricted by file-size limitations, and so FTP continued also to be a popular route.
- **More frequent updates of the dataset:** Although conceived as a climatological product, users were keen to have data up until the present day for their analysis. Users were referred to the Copernicus Climate Change Service (C3S) Climate Data Store (https://cds-dev.copernicus-climate.eu/cdsapp#!/dataset/satellite-ocean-colour?tab=overview) for Chl-a and Rrs that has more frequent updates.
- **Companion datasets such as Sea Surface Temperature (SST):** User were referred to the broader CCI set of products
- **Higher or lower spatial resolution depending on application:** higher resolution for coastal/local studies and lower resolution, to reduce file size, for modelling applications.

3.4.5. CCI+ Workshops

In addition to the 2025 survey, three online workshops were held with participants to showcase the survey findings and seek more detailed feedback. A detailed summary is provided as an Appendix to Kulk and Martinez-Vicente (2025), and this section summarises relevant elements for the OC-CCI project.

Themes from the feedback included:

- **Improved communication of documentation** – users don't seem to find it and then make assumptions… uncertainty & bias are not well understood alongside the impacts/how we reproject the data. In some cases, users are not finding/reading the Product User Guide, whereas for others, shorter webpage-based FAQ-type answers may be easier to digest.

- **Uncertainty and bias in coastal waters** - how to both reduce and improve communication of artefacts to be aware of – change in bias when MERIS ends, and performance of the atmospheric correction and chlorophyll algorithms. We are likely to be underestimating uncertainty in coastal waters, as we don't correct for adjacency, and the algorithms aren't suited to all waters – high latitudes and Amazon outflow were mentioned in particular.

- **Improved resolution**:
  - EO users would like higher spatial resolution for all products, i.e. move to 1 km, while modellers are generally happier with a coarser resolution.
  - Having more wavelength bands and potential for users to generate data with adjusted wavebands to suit their needs – some algorithms/models require specific wavebands, and so when no available the OC-CCI data cannot be used as an input.

- **Additional products** - In particular, PAR was strongly requested.

4. Discussion & Conclusions

The Phase 1 consultation meeting and the user survey revealed that requirements vary depending on model use. For example, the spatial resolution requested by those using global models was 1 degree, whereas those using regional models required spatial resolutions ranging from 4-25 km. The needed temporal resolution was one month for global models and between one day and one week for regional models. The required timeliness of product delivery was typically not crucial in research models, whereas products were required within 1 day of data acquisition in operational models. Error specification was considered crucial by practically all the users. Error reduction was deemed desirable, but there was insufficient clarity regarding what might be feasible in the near-term (3–5 years).

In the OC CCI* phase 2 survey, the minimum requirement for horizontal resolution is now 1-10 km, and the goal is 0.1-1 km – this aligns with the requirements expressed in the Phase 2 user survey in 2015; see **Figure 3-5**. Regarding temporal resolution, the threshold has also decreased, with a preference for 1 week, whereas the goal is 1 day. Again, compared to the 2015 survey, this request was already being seen; see **Figure 3-5**.

Users were also surveyed on the error specification regarding accuracy, precision and stability. The majority of both EO and modelling users saw the benefit, although the percentage was lower for modellers; see **Figure 3-8**. In 2011 and 2015, the majority of users also viewed error characterisation as beneficial.

The review at the start of Phase 2 revealed that many of the wishes expressed by the CMUG and users had been or were already being addressed. Other areas, such as improving the understanding of uncertainties, needed to be addressed across the CCI projects. However, free-form comments highlighted misconceptions and requests for further information, which were addressed where possible. During Phase 3, user feedback via the help desk asked questions similar to those requested in previous phases. Users remain interested inÇ downloading subsets, more frequent updates, companion datasets, and higher- or lower-resolution versions, depending on the application.

Regarding data access, the preferred data format remains NetCDF, with access routes favoring both newer cloud-native routes alongside FTP still having a strong usage; see **Figure 3-10**. Sources of the OC-CCI data are primarily the PML portal alongside the C3S portal. In the analysis of previous used feedback, Section 3.4.4, the C3S portal was being popularised for access to the data and users also saw the benefit of FTP alongside cloud-native access routes.

During the CCI+ Workshops, several prior themes were raised. Of particular importance was further improvement in the communication of how the data is processed and what the products contain. Arefacts/ high uncertainties can be embedded in the time series, and unaware users may interpret them as real phenomena. There was a detailed discussion of how we provide documentation so that it can be found/read, and of the best format for different user communities. Previously, the 2016 CLEO workshop

(Sathyendranath et al., 2016) concluded that a continuous, sustained assessment of product quality was the only means of ensuring that small trends observed at global and regional scales were significant and could be reproduced in historical simulations.

5. Acknowledgements

The OC-CCI Team would like to thank all who responded to the surveys, participated in meetings, or have provided verbal/written feedback.

6. References

Alvain, S, Moulin, C., Dandonneau, Y. and Loisel, H. (2008), Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: A satellite view. *Global Biogeochemical Cycles*, 22, GB3001.

Baird et al. (2016). Remote-sensing reflectance and true colour produced by a coupled hydrodynamic, optical, sediment, biogeochemical model of the Great Barrier Reef, Australia: comparison with 35 satellite data. *Env. Model. Software*, 78, 79-96.

Barton, A.D., Irwin, A.J., Finkel, Z.V., & Stock, C.A. (2016). Anthropogenic climate change drives shift and shuffle in North Atlantic phytoplankton communities. *Proceedings of the National Academy of Sciences*, 113 (11), 2964-2969.

Beaugrand, G., Luczak, C. and Edwards, M. (2009), Rapid biogeographical plankton shifts in the North Atlantic Ocean. *Global Change Biology*, 15(7), 1790-1803.

Behrenfeld, M.J. and Boss, E. (2006), Beam attenuation and chlorophyll concentration as alternative optical indices of phytoplankton biomass. *Journal of Marine Research*, 64, 431- 451.

Behrenfeld, M.J. and P.G. Falkowski (1997), Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology and Oceanography*, 42,1-20.

Behrenfeld, M.J., Boss, E, Siegel, D.A., and Shea, D.M. (2005), Carbon-based ocean productivity and phytoplankton physiology from space. *Global Biogeochemical Cycles*, 19, GB1006, DOI:10.1029/2004JC002527.

Behrenfeld, M.J., R. O'Malley, D.A. Siegel, C. McClain, J. Sarmiento, G. Feldman, A. Milligan, P. Falkowski, R. Letelier, E. Boss, (2006), Climate-driven trends in contemporary ocean productivity. *Nature*, 444, 752-755.

Behrenfeld, M.J., S.C. Doney, I. Lima, E.S. Boss, D.A. Siegel (2013), Physical-ecological interactions of the subarctic Atlantic annual plankton bloom. *Global Biogeochemical Cycles*, 27, 526–540.

Bilir, T.E., Chatterjee, M., Ebi, K.L., Estrada, Y.O., Genova, R.C., Girma, B., Kissel, E.S., Levy, A.N., MacCracken, S., Mastrandrea, P.R., and White, L.L. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1-32

Biogeochemical-Argo Task Team. (2016). The Rationale, Design and Implementation Plan for Biogeochemical-Argo.

Bopp, L., Resplandy, L., Orr, J.C., Doney, S.C., Dunne, J.P., Gehlen, M., & Tjiputra, J. (2013). Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models.

Boyce D.G., Lewis M.R., Worm B. (2010). Global phytoplankton decline over the past century. *Nature*, 466, 591-596.

Bracher, A., M. Vountas, T. Dinter, J.P. Burrows, R. Röttgers, and I. Peeken (2009), Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences*, 6, 751-764.

Brewin, R.J., W.,G. Dall'Olmo, S. Sathyendranath, and N.J. Hardman-Mountford (2012), Particle backscattering as a function of chlorophyll and phytoplankton size structure in the open ocean. *Optics Express*, 20(16), 17632-17652.

Bricaud, A., A. M. Ciotti, and B. Gentili (2012), Spatial-temporal variations in phytoplankton size and colored detrital matter absorption at global and regional scales, as derived from twelve years of SeaWiFS data (1998–2009). *Global Biogeochemical Cycles*, 26, GB1010.

Buitenhuis, E., Vogt, M., Moriarty, R., Bednarsek, N., Doney, S. C., Leblanc, K., et al. & Peloquin, J. (2013). MAREDAT: towards a world atlas of MARine Ecosystem DATa. *Earth System Science Data*, 5, 227-239.

Ciavatta, S., Kay, S., Saux-Picart, S., Butenschön, M. and Allen, J.I. (2016). Decadal reanalysis of biogeochemical indicators and fluxes in the North West European shelf-sea ecosystem. *Journal of Geophysical Research: Oceans*, 121, 1824-1845.

Ciotti, A. M., and A. Bricaud (2006), Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf region off Brazil. *Limnology and Oceanography Methods*, 4, 237–253.

Claustre, H., Legendre, L., Boyd, P. W., & Levy, M. (2021). The Oceans' Biological Carbon Pumps: Framework for a Research Observational Community Approach. *Frontiers in Marine Science*, 8(November), 1–21. https://doi.org/10.3389/fmars.2021.780052

CMUG. (2010). Requirement baseline document. Version 1.2. August 2010.

CMUG. (2021). Climate Modelling User Group Quality Assessment Report (Issue 2.1).

Dall'Olmo, G., and T. K. Westberry (2009). Significant contribution of large particles to optical backscattering in the open ocean. *Biogeosciences*, 6(6), 947–967.

De Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., et al. & Carmichael, M. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science*, 348(6237), 1261605.

Dutkiewicz S, Hickman A.E., Jahn O., Gregg W.W., Mouw C.B., Follows M.J. (2015). Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model. *Biogeoscience*, 12, 4447-4481, DOI:10.5194/bg-12-4447-2015

Ford, D. (2020). Assessing the role and consistency of satellite observation products in global physical–biogeochemical ocean reanalysis, *Ocean Sci.*, 16, 875–893, https://doi.org/10.5194/os-16-875-2020

Fujii M., Boss E., Chai F. (2007). The value of adding optics to ecosystem models: a case study. *Biogeosciences Discuss.*, 4, 1585-1631.

GCOS. (2004). Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC (GCOS-92, WMO/TD-No 1219). Available at http://www.wmo.ch/pages/prog/gcos/Publications/gcos-92_GIP.pdf . For the satellite supplement see http://www.ceos.org/CEOS%20Response%20to%20the%20GCOS%20IP.pdf and http://www.wmo.int/pages/prog/gcos/Publications/gcos-107.pdf

GCOS. (2011). Systematic Observation Requirements for Satellite-Based Data Products For Climate (GCOS-154). Available online at https://www.wmo.int/pages/prog/gcos/Publications/gcos-154.pdf

GCOS. (2016). The Global Observing System for Climate: Implementation Needs (GCOS-200, GOOS-214). Available online at http://unfccc.int/files/science/workstreams/systematic_observation/application/pdf/gcos_ip_10oct2016.pdf

GCOS. (2022). The 2022 GCOS Implementation Plan (GCOS-244). Available online at https://library.wmo.int/records/item/58104-the-2022-gcos-implementation-plan-gcos-244

Gordon H.R. and Morel A. (1983). *Remote assessment of ocean color for interpretation of satellite visible imagery: A review*, Vol. Springer-Verlag, Heidelberg.

Gregg W.W., Casey N.W. (2010). Improving the consistency of ocean color data: A step toward climate data records, Geophys. *Res. Lett.*, 37, L04605, DOI:10.1029/2009GL041893.

Gruber, N. (2011). Warming up, turning sour, losing breath: ocean biogeochemistry under global change. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 369(1943), 1980-1996.

Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., et. Al. & Coelho, L. P. (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature*.

Hashioka, T., Vogt, M., Yamanaka, Y., Le Quere, C., Buitenhuis, E. T., Aita, M. N., et al. & Sailley, S. F. (2013). Phytoplankton competition during the spring bloom in four plankton functional type models.

Hauck, J., Zeising, M., Le Quéré, C., Gruber, N., Bakker, D. C. E., Bopp, L., Chau, T. T. T., Gürses, Ö., Ilyina, T., Landschützer, P., Lenton, A., Resplandy, L., Rödenbeck, C., Schwinger, J., & Séférian, R. (2020). Consistency and Challenges in the Ocean Carbon Sink Estimate for the Global Carbon Budget. *Frontiers in Marine Science*, 7(October), 1–22. https://doi.org/10.3389/fmars.2020.571720

Henson et al. (2010). Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity. *Biogeosciences*, 7, 621-640.

Henson, S.A., Cole, H.S., Hopkins, J., Martin, A.P., and Yool, A. (2018) Detection of climate change-driven trends in phytoplankton phenology, *Glob Change Biol*. 2018;24:e101–e111.

Henson, S.A., R. Sanders, E. Madsen, P. J. Morris, F. Le Moigne, and G.D. Quartly (2011), A reduced estimate of the strength of the ocean's biological carbon pump, *Geophysical Research Letters*, 38, L04606.

Hirata, T., N.J Hardman-Mountford, R.J.W. Brewin, J. Aiken, R. Barlow, K. Suzuki, T. Isada, E. Howell, T. Hashiokia, M. Noguchi-Aita, and Y. Yamanaka. (2011). Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences*, 8, 311-327.

IPCC. (2014). Summary for policymakers. In: Climate Change 2014: Impacts,Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of *Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea,

Jackson, T. (2020). ESA Ocean Colour Climate Change Initiative – Phase 3 Product User Guide for v5 . 0 Dataset. 1–50.

Jackson, T., Sathyendranath, S., & Mélin, F. (2017). An improved optical classification scheme for the Ocean Colour Essential Climate Variable and its applications. Remote Sensing of Environment, 203, 152–161. https://doi.org/10.1016/j.rse.2017.03.036

Kostadinov, T.S., D.A. Siegel, and S. Maritorena. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *Journal of Geophysical Research*, 114, C09015.

Kostadinov, T.S., D.A. Siegel, and S. Maritorena. (2010), Global variability of phytoplankton functional types from space: Assessment via the particle size distribution. *Biogeosciences*, 7, 3239-3257.

Kulk, G. and Martinez-Vicente, V. (2025) PHYTO-CCI User Requirements Document, D1.1, v1.0

Laufkötter, C., Vogt, M., Gruber, N., Aita-Noguchi, M., Aumont, O., Bopp, L., et al. & Hauck, J. (2015). Drivers and uncertainties of future global marine primary production in marine ecosystem models. Biogeosciences, 12, 6955-6984.

Laufkötter, C., Vogt, M., Gruber, N., Aumont, O., Bopp, L., Doney, S. C., et al. & Seferian, R. (2016). Projected decreases in future marine export production: the role of the carbon flux through the upper ocean ecosystem. *Biogeosciences*, 13, 4023-4047.

Laws, E.A., P.G. Falkowski, W.O. Smith, H. Ducklow, and J.J. McCarthy (2000). Temperature effects on export production in the open ocean. *Global Biogeochemical Cycles*, 14, 1231– 1246.

Le Quere, C. L., Harrison, S. P., Colin Prentice, I., Buitenhuis, E. T., Aumont, O., Bopp, L., et al. & Klaas, C. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology*, 11(11), 2016-2040.

Mishonov, A.V., W.D. Gardner, and M.J. Richardson (2003), Remote sensing and surface POC concentration in the South Atlantic. *Deep-Sea Research Part II*, 50 (22–26), 2997–3015

Mouw, C.B. and J.A. Yoder (2010), Optical determination of phytoplankton size composition from global SeaWiFS imagery. *Journal of Geophysical Research*,115, C12018.

Müller, D., Krasemann, H., Brewin, R.J.W., Brockmann, C., Deschamps, P-Y., Doerffer, R., Fomferra, N., Franz, B.A., Grant, M.G., Groom, S.B., Mélin, F., Platt, T., Regner, P., Sathyendranath, S., Steinmetz, F., Swinton J. (2015). Spatial and Temporal Homogeneity of satellite data retrieval due to systematic effects in Atmospheric Correction Processors, *Remote Sensing of Environment*, 162, 257-260.

Nair, A., S. Sathyendranath, T. Platt, J. Morales, V. Stuart, M.-H. Forget, E. Devred, and H. Bouman (2008), Remote sensing of phytoplankton functional types. *Remote Sensing of Environment*, 112, 3366–3375.

Ohring, O., Wielicki, B., Spencer, R., Emery, B., Datla, R. (2005). Satellite Instrument Calibration for measuring global climate change. Report of a workshop. *Bulletin of American Meteorological Society*, September 2005, 1303-1313.

Poloczanska, E.S., Burrows, M.T., Brown, C.J., Garcia Molinos, J., Halpern, B.S., Hoegh-Guldberg, O., et al. & Sydeman, W.J. (2016). Responses of marine organisms to climate change across oceans. *Frontiers in Marine Science*, 3, 62.

Preiserdorfer R.W. (1961). Application of radiative transfer theory to light measurements in the sea. *Int. Geophys. Geod. Monogr.*, 10:11 - 29

Reid, P.C., Johns, D.G., Edwards, M., Starr, M., Poulin, M., & Snoeijs, P. (2007). A biological consequence of reducing Arctic ice cover: arrival of the Pacific diatom Neodenticula seminae in the North Atlantic for the first time in 800 000 years. *Global Change Biology*, 13(9), 1910-1921.

Rivero-Calle, S., Gnanadesikan, A., Del Castillo, C. E., Balch, W. M., & Guikema, S. D. (2015). Multidecadal increase in North Atlantic coccolithophores and the potential role of rising CO2. Science, 350(6267), 1533-1537.

Robinson, C., Poulton, A.J., Holligan, P.M., Baker, A.R., Forster, G., Gist, N., et al. & Woodward, E.M.S. (2006). The Atlantic Meridional Transect (AMT) programme: a contextual view 1995–2005. *Deep Sea Research Part II: Topical Studies in Oceanography*, 53(14), 1485-1515.

Roy, S., S. Sathyendranath, H. Bouman, and T. Platt. (2013). The global distribution of phytoplankton size spectrum and size classes from their light-absorption spectra derived from satellite data. *Remote Sensing of Environment*, 139, 185–197.

Sailley, S.F., Vogt, M., Doney, S.C., Aita, M.N., Bopp, L., Buitenhuis, E.T., et al. & Yamanaka, Y. (2013). Comparing food web structures and dynamics across a suite of global marine ecosystem models. *Ecological modelling*, 261, 43-57.

Sathyendranath, S., Bracher, A., Brockmann, C., Platt, T., Ramon, D., Regner, P. (2016) Colour and Light in the Ocean (CLEO) 2016: A Scientific Roadmap from the Workshop Organised by ESA and PML. Held at ESRIN, Frascati, Italy on 6-8 September, 2016 https://odnature.naturalsciences.be/downloads/publications/cleo_scientific_roadmap_from_esrin_workshop__6-8_september_2016.pdf

Sathyendranath, S., Brewin, R. J. W., Brockmann, C., Brotas, V., Calton, B., Chuprin, A., Cipollini, P., Couto, A. B., Dingle, J., Doerffer, R., Donlon, C., Dowell, M., Farman, A., Grant, M., Groom, S., Horseman, A., Jackson, T., Krasemann, H., Lavender, S., … Platt, T. (2019). An ocean-colour time series for use in climate studies: The experience of the ocean-colour climate change initiative (OC-CCI). Sensors (Switzerland), 19(19). https://doi.org/10.3390/s19194285

Siegel, D.A., Behrenfeld, M.J., Maritorena, S., McClain, C.R., Antoine, D., Bailey, S.W., Bontempi, P.S., Boss, E.S., Dierssen, H.M., Doney, S.C., Eplee, R.E. Jr., Evans, R.H., Feldman, G.C., Fields. E., Franz, B.A., Kuring, N.A.,Mengelt, C., Nelson, N.B., Patt, F.S., Robinson, W.D., Sarmiento, J.L., Swan, C.M., Werdell, P.J., Westberry, T.K., Wilding, J.G., Yoder, J.A. (2013). Regional to global assessments of phytoplankton dynamics from the SeaWiFS mission. *Remote Sensing of the Environment*, 135, 77-91.

Siegel, D.A., Buesseler, K.O., Behrenfeld, M.J., Benitez-Nelson, C.R., Boss, E., Brzezinski, M.A., et al. & Perry, M.J. (2016). Prediction of the Export and Fate of Global Ocean Net Primary Production: The EXPORTS Science Plan. *Frontiers in Marine Science*, 3, 22.

Siegel, D.A., K.O. Buesseler, S.C. Doney, S.F. Sailley, M.J. Behrenfeld, and P.W. Boyd (2014), Global assessment of ocean carbon export by combining satellite observations and food-web models, *Global Biogeochemical Cycles*, 28, 181–196, doi:10.1002/2013GB004743.

Siegel, D.A., S. Maritorena, N.B. Nelson, and M.J. Behrenfeld. (2005). Independence and interdependences among global ocean color properties Reassessing the bio-optical assumption. *Journal of Geophysical Research*, 110, C07011.

Stramski, D., R.A. Reynolds, M. Kahru, and B.G. Mitchell (1999), Estimation of particulate organic carbon in the ocean from satellite remote sensing. *Science*, 285, 239–242.

Taylor, K.E., Stouffer, R.J., & Meehl, G.A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4), 485.

Van Oostende, M., Hieronymi, M., Krasemann, H., Baschek, B., Röttgers, R. (2022). Correction of inter-mission inconsistencies in merged ocean colour satellite data, Frontiers in Remote Sensing, 3, doi:/10.3389/frsen.2022.882418

Vogt, M., Hashioka, T., Payne, M.R., Buitenhuis, E.T., Quéré, C.L., Alvain, S., et al. & Lima, I. (2013). The distribution, dominance patterns and ecological niches of plankton functional types in Dynamic Green Ocean Models and satellite estimates. *Biogeosciences Discussions*, 10(11), 17193-17247.

Westberry, T.K., M.J. Behrenfeld, D.A. Siegel, E. Boss (2008), Carbon-based primary productivity modeling with vertically resolved photoacclimation. *Global Biogeochemical Cycles*, 22, GB2024.

7. Appendix I: Glossary

	Definitions
Horizontal resolution	the area over which one value of the variable is representative of
Temporal Resolution	the temporal frequency at which the measurements are required.
Accuracy	the measure of the non-random, systematic error, or bias, that defines the offset between the measured value and the true value that constitutes the SI absolute standard
Precision	the measure of reproducibility or repeatability of the measurement without reference to an international standard so that precision is a measure of the random and not the systematic error. Suitable averaging of the random error can improve the precision of the measurement but does not establish the systematic error of the observation.
Stability	a term often invoked with respect to long-term records when no absolute standard is available to quantitatively establish the systematic error – the bias defining the time-dependent (or instrument-dependent) difference between the observed quantity and the true
LEVELS	
Threshold	minimum requirement to be met to ensure that data are useful
Goal	an ideal requirement above which further improvements are not necessary
Breakthrough	an intermediate level between "threshold" and "goal" which, if achieved, would result in a significant improvement for the targeted application. The "breakthrough" level may be considered as an optimum, from cost benefit point of view when planning or designing observing systems (or models)

8. Appendix II: Phase 1, 2011 User Survey and Additional Information

The majority of respondents were from the EO community, and they were mainly interested in global-scale analysis, primary production studies, ecological provinces and fisheries variability (**Figure 8-1**). Other applications of interest were photochemistry, bio-optical studies, Carbon cycle studies and water quality monitoring.

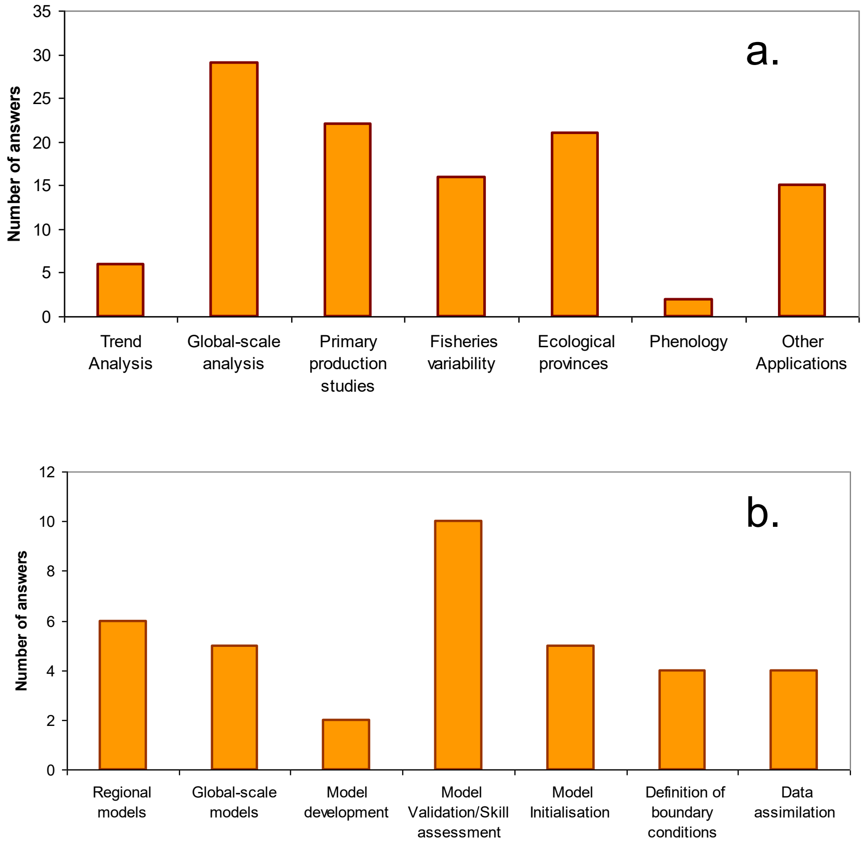


Figure 8-1: Areas of research of a) EO scientists and b) Climate modellers. Note that this was a multiple-choice question; hence the total number of answers to it was greater than the number of responses

8.1. Chlorophyll-a concentration (Chl-a)

The following section summarises the results from the questionnaire concerning the requirements for Chl-a. The total number of responding scientists to the questionnaire was 78. This total sample was separated into Climate modellers (N= 15) and Earth Observation scientists (N= 63).

Concerning the horizontal resolution (**Table 8-1**) required for Chl-a, the requirements from the EO scientists were more stringent than those from GCOS or CMUG. CCI modellers' requirements (Goal and Breakthrough) agreed with CMUG estimates. For this group of users, the important feature was to obtain satellite Chl-a data at a resolution similar to global ocean models (with a grid size of ~10km or coarser). The reasons expressed in the questionnaire for the EO community to request 0.1 – 1 km resolution was to allow for the capture of small-scale processes in coastal studies of climate change and to incorporate sub-mesoscale processes into global models.

The observing cycles (temporal frequency of sampling) required was one week compared to one day from CMUG, but also one week from GCOS. Weekly Chl-a data at a resolution of at least one week is required to account for the limitations imposed by cloud cover. The monthly resolution would be sufficient for model evaluation, but the weekly frequency (at least) was critical for data assimilation. However, any product averaged over longer periods than 1 week would miss the timing of phytoplankton blooms completely and hence would not allow the study of seasonality and changes thereof.

Table 8-1: Spatial Resolution

Ocean chlorophyll	Horizontal Resolution			Observing Cycle		
Source of requirement	Goal	Breakthrough	Threshold	Goal	Breakthrough	Threshold
Requirement from GCOS (2011 and 2016)	30 km originally, reduced to 4 km			7 d		
Requirement from CMUG		4 km			1 d	
Requirement from CCI – Modellers	0.1 - 1 km	1 - 10 km	1 - 10 km	1 d	7 d	30 d
Requirement from CCI – EO scientists	0.1 - 1 km	0.1 -1 km	1 - 10 km	1 d	7 d	30 d

Breakthrough accuracy requirement levels from CMUG were similar to those that emerge from our questionnaire and were more stringent than the original GCOS (2011) breakthrough requirements. However, they were in-line with the revised GCOS (2016) requirements. Modellers require an accuracy of 10 – 25 % for the precision of Chl-a (note that this was at variance with the view of many of the experts who attended the London meeting, who considered it was more important to have high accuracy than high precision). Their request was based on the needs of unbiased validation data sets and data sets for assimilation, as well as their awareness of the capabilities of satellites. For EO scientists, the accuracy requirement was 10 – 25 %. A reason for this threshold expressed in the survey was to avoid that the bias propagates into derived variables such as primary production or gas-flux computations.

Precision requirements have not been explicitly recorded in the GCOS or CMUG reports, but the expectation of modellers and EO scientists was likely to be a similar precision to in-situ measurements of Chl-a (~ 10%).

Table 8-2: Accuracy and Precision

Ocean chlorophyll	Accuracy			Precision		
Source of requirement	Goal	Breakthrough	Threshold	Goal	Breakthrough	Threshold
Requirement from GCOS (2011 and 2016)*	30 %					
Requirement from CMUG		30%			30%	
Requirement from CCI - Modellers	< 10 %	10 – 25 %	10 – 25 %	< 10 %	< 10 %	10 - 25 %
Requirement from CCI	< 10	10 – 25 %	10 – 25	< 10	10 - 25 %	10 - 25 %

– EO scientists	%		%	%	

* GOCS 'accuracy' estimates are "indicative of acceptable overall levels for the uncertainties of product values" (GCOS, 2011).

Stability of the monitoring of the ECV is a characteristic that it is crucial for climate change studies; therefore, it was included in the CCI Ocean Colour questionnaire. Results show a similar request from modellers and EO scientists. For modellers, 10 % variation over 10 years would be considered sufficient for data assimilation studies, whilst a 5 % variation over 10 years would be a maximum to allow for climate studies. EO scientists consider that given recent publications (e.g. Boyce et al. 2010), where a 1 % annual decline in Chl-a is observed, a stability of at least 1 % over ten years would be required for detecting climate change.

Table 8-3: Stability

Ocean chlorophyll	Stability (over 10 years)		
Source of requirement	Goal	Breakthrough	Threshold
Requirement from GCOS (2011 and 2016)		3%	
Requirement from CMUG			
Requirement from CCI - Modellers	1%	1%	10%
Requirement from CCI – EO scientists	1%	1-2 %	5%

The users were also interrogated about the importance of having error estimates assigned to Chl-a from EO. There was a very clear requirement from both communities requesting product error estimates to be made available along with the product values (100% of modellers and 95% of EO scientists believe having error estimates was important).

8.2. Normalised Water-Leaving radiance (nLw)

Water-leaving radiance is the upwelling radiance just below the sea surface, which is a function of the view angle. Normalised water-leaving radiance approximates the radiance that would leave the sea surface in the absence of an atmosphere, and with the sun at the zenith.

For EO scientists, what came across the questionnaire is, in the words of one of the users: "*We need nLw values for validation calibration studies and development of empirical algorithms for various products. No having nLw will stop consequent improvement of the local empirical algorithms - global and semi-analytical algorithm usually fail in coastal areas.*" So nLw was required mainly to have the possibility to develop and improve bio-optical algorithms. Modellers did not use nLw directly in their coupled models. Further, none of the modellers were using nLw to compare with model outputs either at the water surface or extrapolated through the atmosphere. This was highlighted in the CCI collocation meeting (September 2010) as a potential source of development for the Climate models. Therefore, a specific question on it was included in the questionnaire. The negative response indicated that the modelling community was not yet ready to ingest the ocean colour imagery at the top of the atmosphere (as observed by the satellite).

Table 8-4: Threshold (minimum) requirements for water-leaving radiance (nLw)

Water Leaving radiance - nLw	Horizontal Resolution	Observing Cycle	Accuracy	Precision	Stability (over 10yr)

Source of requirement	Threshold	Threshold	Threshold	Threshold	Threshold
Requirement from GCOS (2011 and 2016)	4 km	1 d	5 %*		0.5 %
Requirement from CCI – Modellers	10 - 100 km	30 d	10 - 25 %	~ 25 %	2 %
Requirement from CCI - EO scientists	0.1 - 1 km	1 d	~ 10 %	~ 10 %	1 %

*this 5% requirement was specifically for the blue and green wavelengths.

8.3. Spectral attenuation coefficient for downwelling irradiance (Kd)

The attenuation coefficient for downwelling irradiance provides the rate of decrease of the irradiance with unit distance, and per unit incident flux.

EO scientists consider this parameter important: "*Kd is a core underwater light climate variable that is linked to primary productivity; bathymetry assessment; light penetration to bottom estimates etc... It also functions as the only Quasi AOP that is measurable with uncalibrated instruments – it is a core variable for achieving optical closure between in situ measured IOP's and AOPs*". For modellers: "*Kd could be naturally assimilated too but is not used as of today*", because '*May be used to provide information about light availability, rather than modelling light penetration as a function of chlorophyll. Requires at least similar accuracy to be worth using for Case 1 water*".

Table 8-5: Threshold (minimum) requirements for spectral attenuation coefficient for downwelling irradiance (Kd)

Kd	Horizontal Resolution	Observing Cycle	Accuracy	Precision	Stability (over 10 yr)
Source of requirement	Threshold	Threshold	Threshold	Threshold	Threshold
Requirement from CCI – Modellers	1 - 10 km	7 d	10 - 25 %	10 - 25 %	2 %
Requirement from CCI - EO scientists	0.1 - 1 km	1 d	~ 10 %	~10 %	1 %

8.4. Inherent optical properties (IOPs)

Inherent optical properties (IOPs) include absorption (*a*), scattering (*b*) and backscattering (b_b) and were defined as optical properties independent of variations in the angular distribution of the incident light (Preisendorfer, 1976).

Some EO scientists consider IOP as the most important set of variables to be retrieved from ocean colour: "*IOPs are true physical properties of the ocean.[…]I would be happy to have just these products from ocean color.*" Mainly they are: "*Needed as input for biogeochemical models (absorption by phytoplankton and CDOM, backscattering etc'). […] Don't think the products are in the accuracy range I would have hoped...*".

Table 8-6: Threshold (minimum) requirements for Inherent Optical Properties (IOP)

IOP	Horizontal Resolution	Observing Cycle	Accuracy	Precision	Stability (over 10 yr)
Source of requirement	Threshold	Threshold	Threshold	Threshold	Threshold
Requirement from CCI – Modellers	1 - 10 km	1 to 7 d	10 - 25 %	10 - 25 %	~ 2 %
Requirement from CCI - EO scientists	0.1 - 1 km	1 d	~ 10 %	~ 10 %	~ 2%

8.5. Coloured dissolved organic matter (CDOM)

Absorption by CDOM is an IOP, as it is the absorption produced by the dissolved (organic) matter in the water (expressed in m^{-1}). This variable can be approximated indirectly from Chl-a concentration in areas of the ocean where there are no external sources of dissolved matter (for instance rivers). In this context, EO scientists recognised the importance of CDOM estimates as a proxy for riverine inputs, and therefore CDOM would be useful for quantifying the effects of climate change on river discharge. Although modellers did not use CDOM measurements, they were aware of the potential importance for validation of models outputs.

Table 8-7: Threshold (minimum) requirements for Coloured Dissolved Organic Matter (CDOM)

CDOM	Horizontal Resolution	Observing Cycle	Accuracy	Precision	Stability (over 10 yr)
Source of requirement	Threshold	Threshold	Threshold	Threshold	Threshold
Requirement from CCI – Modellers	1 - 10 km	7 d	25 - 50 %	10 - 25 %	10%
Requirement from CCI - EO scientists	0.1 - 1 km	1 d	~ 10 %	~ 10 %	5%

8.6. Suspended particulate matter (SPM)

An operational definition of TSM is "the amount of the material retained on a GF/F filter of 0.47µm pore size by unit volume", which pools together all kinds of suspended particles of different origin, sizes, shapes, internal composition and optical properties. It was recognised across the responses of both groups in the questionnaire as a very important variable for monitoring the climate change in coastal areas, although not so much in oceanic waters, where concentrations are low and mainly related to Chl-a variations.

Table 8-8: Threshold (minimum) requirements for Suspended Particulate Matter (SPM)

SPM	Horizontal Resolution	Observing Cycle	Accuracy	Precision	Stability (over 10 yr)
Source of requirement	Threshold	Threshold	Threshold	Threshold	Threshold
Requirement from CCI - Modellers	1 - 10 km	1 - 7 d	10 - 25 %	10 - 25 %	5%
Requirement from CCI - EO scientists	0.1 - 1 km	1 d	10 - 25 %	~ 10 %	2%

8.7. Phytoplankton functional types (PFT)

The algorithms that separate the ocean colour signal among different phytoplankton types. At the time, it was mainly an experimental product that, in most cases, considered phytoplankton functional types in oceanic waters, and it referred often to separation by size class. Both groups of users agreed on the importance of PFT for ecosystem model validation and that the spatial requirements can be relaxed in this case. One modeller synthetises the view in the comments: “*We need phytoplankton functional types in order to study the interaction of phytoplankton with the carbon cycle, to detect changes in biodiversity/biogeochemical provinces, study the feedbacks between plankton productivity and fisheries etcetc. We desperately need PFT measurements to validate our state-of-the-art models, in particular since we do not yet know the complexity needed to address our biological and biogeochemical questions. Optimally, we'd have phytoplankton-specific HPLC pigments also from space, but I think this is beyond what this ESA mission can do for us. The consequences of not having any observationally-determined PFTs would be dire - we wouldn't be able to validate our models. We are already using Severine Alvain's PFTs from space to validate the models, but we use dominance and frequencies rather than monthly mean absolute contribution to total chlorophyll.”*

Table 8-9: Threshold (minimum) requirements for Phytoplankton Functional Types (PFT)

PFT	Horizontal Resolution	Observing Cycle	Accuracy	Precision	Stability (over 10yr)
Source of requirement	Threshold	Threshold	Threshold	Threshold	Threshold
Requirement from CCI – Modellers	1 - 10 km	7 d	10 - 25 %	10 - 25 %	2%
Requirement from CCI - EO scientists	1 - 10 km	7 d	10 - 25 %	10 - 25 %	5%

8.8. Particle size distribution (PSD)

At the time, this was another experimental product. Although both users groups were aware of the importance to retrieve this variable from space, the application into climate models had not been yet explored.

Table 8-10: Threshold (minimum) requirements for Particle Size Distribution (PSD)

PSD	Horizontal Resolution	Observing Cycle	Accuracy	Precision	Stability (over 10 yr)
Source of requirement	Threshold	Threshold	Threshold	Threshold	Threshold
Requirement from CCI - Modellers	1 - 10 km	7 d	10 - 25 %	10 - 25 %	2%
Requirement from CCI - EO scientists	0.1 - 1 km	7 d	10 - 25 %	~ 25 %	5%

8.9. Use of CZCS in climate studies

The Coastal Zone Color Scanner (CZCS) measured reflected solar energy in six channels, at a spatial resolution of 800 meters. CZCS lay the foundations for subsequent satellite ocean colour sensors, and formed a cornerstone for international efforts to understand the ocean's role in the carbon cycle. It was recognised by the respondents of the questionnaire that the CZCS data may be useful to extend studies of climate change using ocean colour backwards in time. However, a large proportion of the users did not take CZCS as an important source of climate-related data (Fig. 3-5). Some of the reasons provided by the users in this respect are: "*Poorly characterised data, cannot reliably estimate long term variability or trends when comparing with SeaWiFS/MODIS*" and "*Uncertainties in the processing of the data sets and linking to the modern data sets. Limited number of bands for inversions. Poor sampling.* " . Overall, according to users, CZCS could be a useful tool for phenology or fisheries studies if careful processing is made and it is not "mixed" with data from other sensors. The International Ocean Colour Coordinating Group (IOCCG) review noted that responses on the utility of CZCS data may have been subtly different with a slight re-phrasing of the question on this sensor.

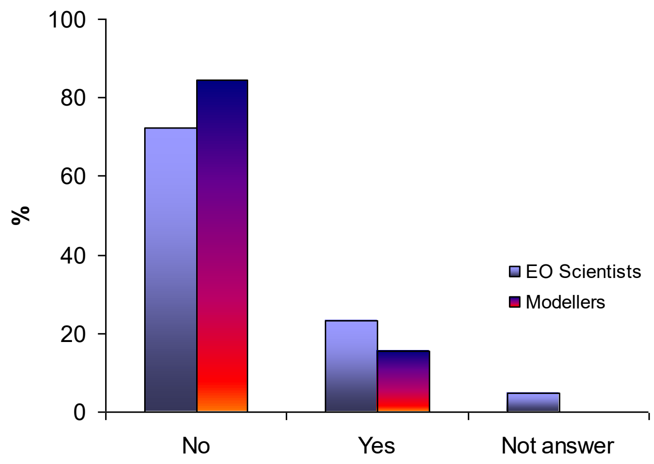


Figure 8-2: Percentage of use of CZCS data separated by user type.

8.10. Requirements for data manipulation and other tools

This section of the questionnaire was dedicated to obtaining an indication of the requirements of the OC users concerning tools for analysis of remote sensing data. The first part of this section asked about the tools provided by third parties being used and requested details on the specific third party tools. From the answers in the questionnaire, it can be seen (Fig 3.6) that there was a diversity of applications for the tools in use. In particular, modellers tend to use Matlab and R as a primary tool for these applications. EO scientists use, apart from Matlab and R, mainly SeaDAS, BEAM, IDL, Python and ERDAS, as well as SISCAL and ARC-GIS tools.

Some capabilities of the tools considered particularly useful by modellers are: "*point extraction*", "*can tailor things to my needs", "Netcdf library","Regridding", "flexibility", "easy reading of the data".* Some capabilities and desirable features of the tools used by EO scientists were (in their own words):

- *Scripting*
- *Batch processing - people are too infatuated with GUI's*
- *Statistical analysis and rendering.*
- *Versatility: in some tools, the possibility of having open source access to create my own scripts, etc...*
- *In SeaDAS I do not like the fact that we have a black square where we move the cursor or mouse. In SeaPAK (old program) we worked always directly on the image. I still think that SeaPAK was a great software.*
- *With BEAM tool (from ESA) I also think that although has really great potential, it is still hard to acquire good knowledge of this tool on our own. I think it would be great a better support for tool learning.*
- *Matlab is great for imagery processing*
- *SISCAL: simple and hassle-free access to marine EO data products!*
- *Can extract data of a single site easily, and can see the temporal and spatial trends*
- *Inter-/extrapolation*
- *Calculation with spectral bands (to try regional algorithms)*
- *ENVI capabilities of display data and read and histogram the data, and some math procedures. It is especially useful to get a spectrum of one pixel in hyperspectral data. It is restricted in the sense that if you want to visualize a chunk of data in 3D, it cannot be done without IDL programming.*
- *SEADAS is a very useful program for visualizing ocean color imagery, and for displaying some small data sets. It has been extremely useful in my work because I can easily convert an HDF file to a binary file with just the click of a button. ENVI can also do that, but it creates a header and other issues.*
- *Matlab for analyses is a much better tool, but cumbersome in the sense that you need to do a lot of programming.*
- *NRL-APS is a very powerful program, based on scripts. It is very easy to use if you have a computer science degree. It takes a long time to learn if you are not.*
- *Great for analysis purposes, easy to work with many different data types.*
- *ease of programming*
- *Comprehensive, efficient and transparent (i.e., with source codes and help) package*
- *Versatility*
- *Integrated with data collection. Able to distribute products to various sites and ftp locations automatically.*
- *Reading L3 processors*
- *Ease of use and easy to learn*
- *Array processing, visualisation and statistics.*

- *Good virtual memory to handle large datasets*
- *If they are flexible in their use and come with a good manual.*
- *Very easy access to NRT and archived RS data with machine friendly interface.*

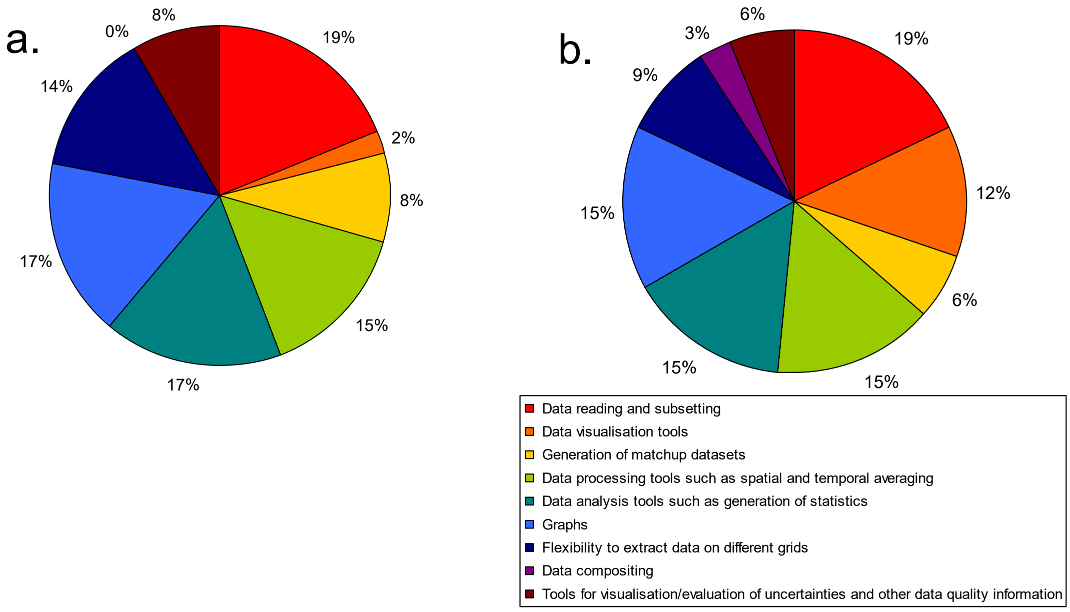


Figure 8-3: Tools used for dealing with Ocean colour products by: a) EO scientists and b) Climate modellers.

In the second part of this section, the users were asked about the tools provided by third parties that they would like this initiative to provide, as well as which specific third party tools they would prefer. The answers in this part show an increased interest in tools for visualisation/evaluation of uncertainties and other data quality information (Fig.3.7), compared with what they use (Fig. 3.6).

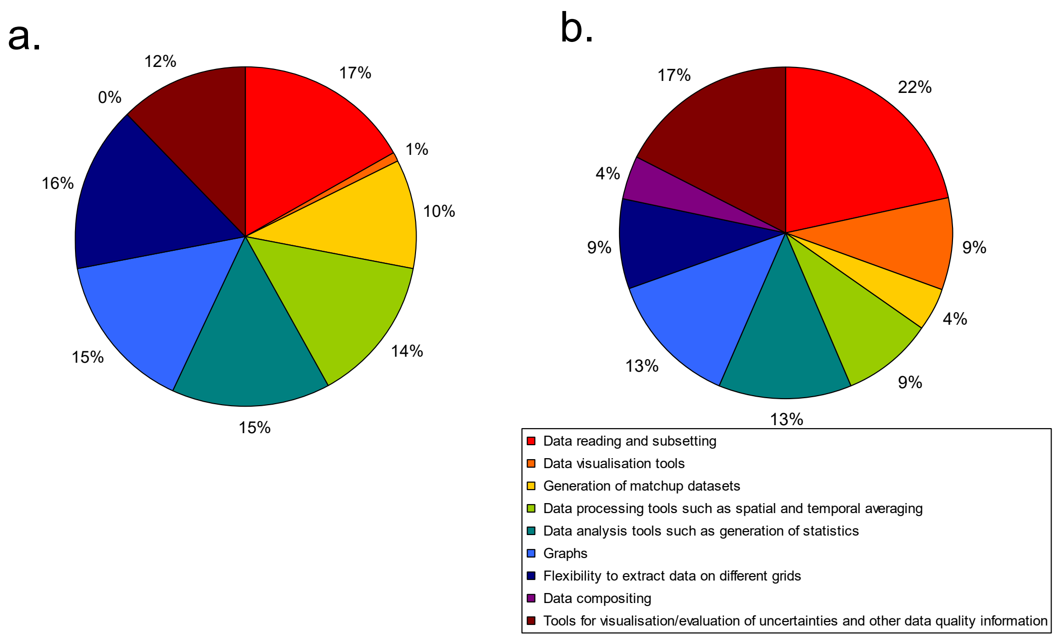


Figure 8-4: Applications that would be used for Ocean colour products (if available) by: a) EO scientists and b) Climate modellers.

Example of some additional capabilities/features that would be particularly useful in a tool are:

1. *Capability to process data from Level 1b to Lt, and get a geometrically corrected Lt file.*
2. *Capability to correct for atmosphere, and produce Lw also in geometrically corrected format.*

The inputs to the file, should be the pixel/line or lat/lon of each of the four corners or simply the lat/lon of the center pixel.

The majority of the users, both modellers and EO scientists, would preferred using free tools that are developed under an open-source license and that came with source-code. However, very few modellers (23%) were aware of the open-source ESA BEAM Toolbox (http://www.brockmann-consult.de/beam/). This was in contrast with the EO scientists that were aware and use (or intend to use) BEAM (97%). Most of the users of BEAM would find it beneficial if all or some of the tools and services above were realised in the BEAM toolbox.

Finally, users were asked if they would like to incorporate analysis, visualisation and processing in their own software programs by using a dedicated library which this initiative might provide. If yes, they were asked to state the preferred programming language for which an API (application programming interface) should be provided, if possible. Results to this question (Fig. 3.8) indicate that the majority of modellers preferred an API for Python followed by Matlab and IDL. EO scientists were more favourable to Matlab, then IDL and finally Python and C++ at a similar level. Other languages such as C, R, ArcGIS, Java, Fortran or Octave did not reach 10% of the responses.

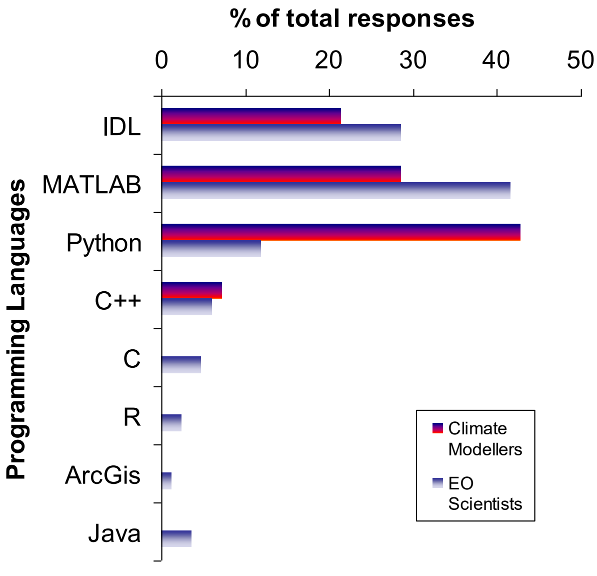


Figure 8-5: Programming languages of preference for Climate modellers and EO Scientists for which an API should be preferentially provided. Because this question was multiple choice, the total percentage can add more than 100.

8.11. Requirements for data formats and access

This part of the questionnaire was dedicated to surveying the preferences of users with respect to data format and access. Results of this section are summarized hereafter. Users were inquired about the level of processing required for their data, and were given the following options (multiple choice):

- Level 1 Geophysical Measurements (e.g. radiances),
- Level 2 Datasets (derived products on original space view),
- Level 3 (e.g. daily, monthly means gridded).

The results indicated significant differences among both users communities (Fig. 3.9 a). The majority of EO scientists preferred to obtain Level 1 data so that they can test different algorithms (including atmospheric corrections and in-water algorithms). In contrast, modellers preferred to obtain Level 3 data that can be easily extracted to be compared with models outputs or included in data assimilation schemes.

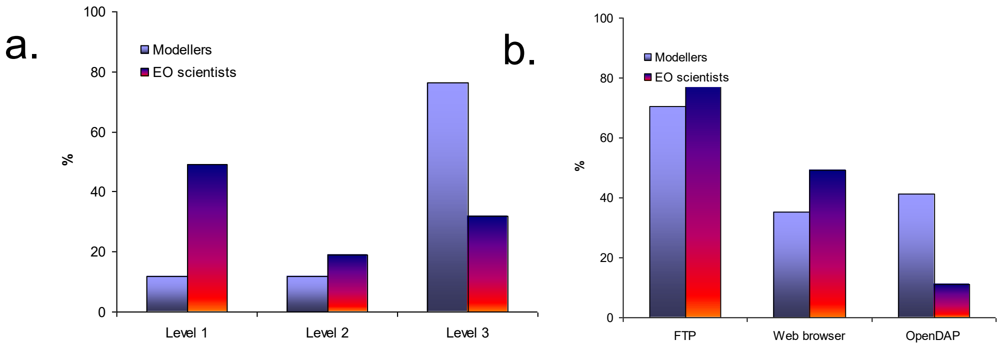


Figure 8-6: Preference of users concerning: a) Level of processing of data and b) Data access. Note that this was a multiple-choice question; hence the total number of answers to it was greater than the number of responses

Concerning the preferred means for accessing the data, differences among users were less important (Fig. 3.9 b). Most of users in both groups coincided in preferring FTP access as a first choice. The preferences then were different for other methods of access: whilst EO scientists preferred Web browser access, modellers preferred OPeNDAP. An example of a comment from a user, concerning data access can illustrate a feeling from users: *"FREE and EASY access to VARIOUS RS data is essential for strengthening application of satellite observation. Additional on-line services or free software might be also helpfull, but access to original RS data is much more important. Fortunately now ESA is more motivated to share data, however there are still a lot of lessons to be learned from Americans."*

There was an overall agreement in preference of CF Compliant NetCDF, and some EO scientists have also requested data in HDF format. Finally, modellers find that having the data in CIMP-5 format would be useful.

9. Appendix III: Phase 2, 2015 User Survey and Additional Information

9.1. Product requirements

From the Phase 1 survey, the following sections summarise the results in terms of (horizontal) spatial resolution, accuracy, precision and stability for both Chl-a and other variables derived from Ocean Colour data: nLw, Kd, IOP, CDOM, SPM, PFT and PSD.

Within the Phase 2 survey, the focus was on the contents of the released OC-CCI dataset, and so the questions asked related to the dataset products: Rrs, IOP, Chl-a and Kd. Of the users sampled, 94 % use satellite-derived ocean-colour data and 93 % were using or interested in using the OC-CCI products; with 27 % using the data to support modelling activities. The actual usage of individual product types was fairly even distributed, but slightly higher for Chl-a; **Figure 9-1**.

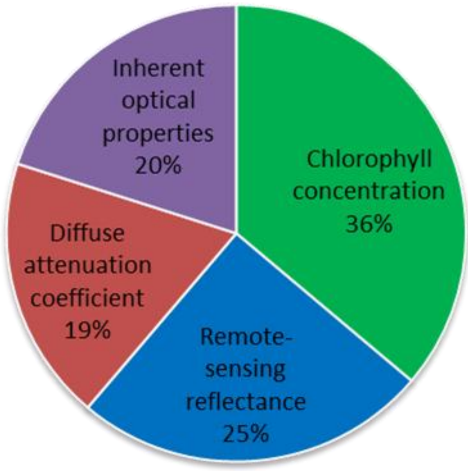


Figure 9-1: Usage of the different product categories.

9.2. Spatial and temporal resolution

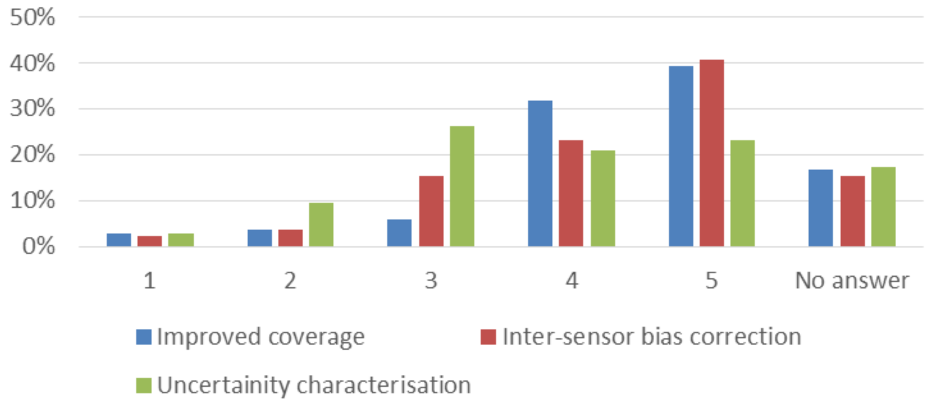
A significant proportion of the users (70 %) would preferred data that had a spatial resolution of between 0.1 and 1 km, which was finer than the standard OC-CCI products. However, the resolution of the products (4 km) does meet the majority's minimum requirement of 1 to 10 km resolution (**Figure 9-2 top**). For the temporal resolution, the preference was for 1 image per day, but a coarser resolution of 1 image per week was an acceptable minimum requirement for the majority of the participants (82 %).

These findings were comparable to the Phase 1 survey where for Chl-a, where the important feature was to obtain satellite Chl-a data at a resolution similar to global ocean models (specified as a grid size of ~10 km or coarser). For the EO community, 0.1 – 1 km resolution allows for the capture of small-scale processes in coastal studies of climate change, and the incorporation of sub-mesoscale processes into global models. The requirement for temporal frequency sampling was one week, as this can account for the limitations imposed by cloud cover. However, having a higher temporal resolution allows transitory features to be captures; becomes more critical for coastal applications.

Figure 9-2: Minimum and ideal requirements in terms of (top) spatial resolution and (bottom) temporal resolution

9.3. Characteristics of the OC- CCI datasets

All three OC-CCI data characteristics were considered important (**Figure 9-3 top**), with an acceptable level of uncertainty being 25 % or less for all the OC-CCI products (**Figure 9-3 bottom**). Also, most of the people responding to the question (75 % of people) feel having uncertainties are indispensable or desirable, while only 15 % class then as not being required; **Figure 9-4**.



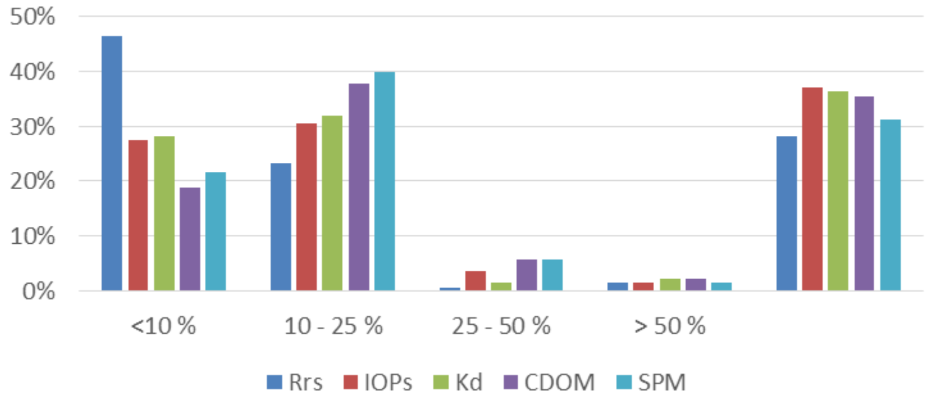


Figure 9-3: The (top) importance of the data characteristics and (bottom) the requirements in terms of an acceptable level of uncertainty.

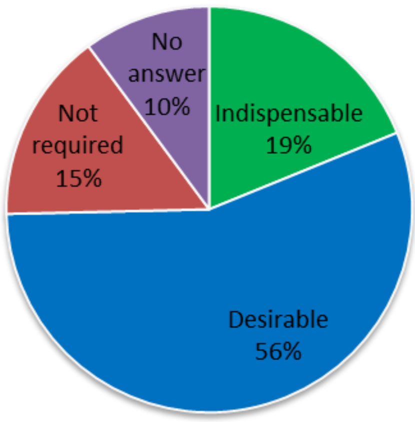


Figure 9-4: Utility of OC-CCI uncertainty estimates

Responses to the question 'Is there any characteristic that is not in the list above that you find important in the OC-CCI products?' include:

- Products and their characteristics:
  - *beam atten coef & Optical backscatter*
  - *PAR, aph and aCDM*
  - *Interpretation of scattering/absorption for carbon biomass.*
  - *temporally up-to-date*
  - *Validity in optical complex waters*
  - *We are using surface reflectance's from MODIS in our data assimilation system. When we compare the reflectances derived from global atmospheric correction algorithms, to those from regionally specific algorithms, there can be substantial differences. Therefore, improving the atmospheric correction algorithms will help reduce the uncertainty in the reflectance products.*
  - *I would like to be able to know if land or bottom reflectance is an issue in the regions I am obtaining data from. Was the land bias removed? How confidence can I be that the signal is water column only and not influenced by bottom?*
- Documentation:
  - *1. Careful documentation of how uncertainty is calculated... and 2. Per-pixel metadata that document exceptions in sensor status, data processing, e.g. NASA L2 flags, perhaps on steroids."*
  - *Documentation on the underlying scientific approach / processing*

- *Some easily accessible assessment of inter-sensor bias n a per pixel basis would be good. For example, how much did the sensors disagree at a given pixel? How much of that is due to noise vs. consistent bias, if that can be quantified?*
- *How do I combine data with SeaWiFS, MODIS or the old CZCS? What are the inter sensor biases?*
- *Band shifting and Large algorithm intercomparison exercise*
- *It is important to keep file sizes to a minimum, whilst maintaining data quality and ease of use - I appreciate this is a compromise.*
- *Readily available*

9.4. Accessing the OC-CCI datasets

Almost half (48 %) of the people responding to this question still download data by FTP (**Figure 9-5**), with HTTP being the next largest method (24 %) and then the Web GIS Portal (12 %).

Approaches that allow on-the-fly sub-setting of the NetCDF files (OPeNDAP and THREDDS) accounted for 15 % together. An issue for the usage of THREDDS was that downloads of greater than 4 GB as a NetCDF file were not possible via the oceancolour.org website; due to the default output format being NetCDF-3 rather than NetCDF-4. This becomes an increasing issue as the full v2.0 OC-CCI dataset volume was 46 TB, compared to 27 TB for v1.0. As an alternative, OPeNDAP allows a binary or ASCII output format of any size.

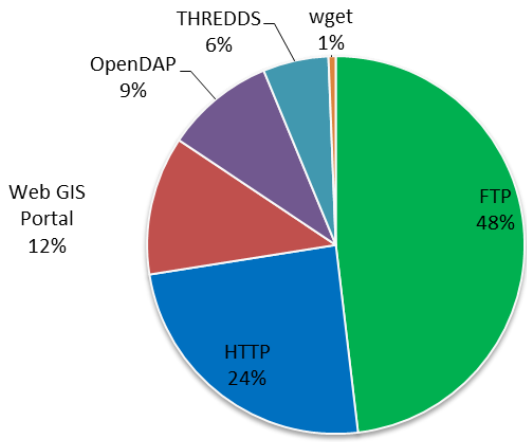


Figure 9-5: Methods for OC-CCI data access

9.5. Help to use the OC- CCI datasets

When asked ‘What feature would make it easier for you to use OC-CCI products?’, and the responses included:

- The download interface:
  - *Access through Giovanni or a Giovanni like interface (3 people) e.g. ‘Available on line with an easy interface. Best if consistent with NASA infrastructure’.*
  - *"Currently I get all my ocean colour data from NASA. I would like to use the ESA system but the front end is not very easy to navigate. As a native english speaker I still struggle with ""limiting the selection to a specific predetermined tag associated with a variable such as its type being chlorophyll"". What does this mean?*
  - *The AOI selection in the composite browser.*

  - *Extraction of time series from selected points or regions, including uncertainties and statistics*
  - *I would like to be able to use shape files rather than just rectangular shapes to define the sample area.*
  - *Easier data access*
  - *Download should be fast.*
  - *Quicklooks should be on the web.*
- Help reading the data:
  - *Scripts in several languages that read in the files. I don't need that, but many users would, I imagine.*
  - *Simpler gridding/Codes to convert to Matlab*
- Product characteristics:
  - Spatial resolution:
    - *1 km normalized water leaving radiance at all visible/NIR channels*
    - *Midlevel (50km or better) resolution weekly filled product over entire period.*
  - Splitting up the products:
    - *Satisfied with the current distribution system. But I am aware that some find the size of the files daunting.*
    - *Do not package all products from a single day/week/month into a single enormous netcdf file - it takes forever to download, is impossible to store and infuriating when you just want chlorophyll*
    - *Modularity in downloading products (i.e., to be able to download only the products required).*
    - *Netcdf files of variables clearly labeled (by filename is file if the convention is given in a document in the same location as the files) on a server so that programmers can: 1. download a zipped file 2. Unzip 3. use netcdf tools of desired language to read data set directly into workspace (e.g. ncread for matlab); then analyse to their heart's content.*
  - *The availability of single-sensor along-track products in addition to merged daily products would allow for extra experiments to be performed.*
  - *Better mapping - there is "salt and pepper" in the data during the SeaWiFS era*
- Documentation:
  - *I like to know something about the validation. Some provenance would be nice.....what methods and what data sets underpin these numbers?*
  - *Global validation especially in Southern Hemisphere*
  - *I would like to know what all the variables are (what is water class 9, are these from Jerlov's old 1968 classifications? there are a huge number of water classification systems out there.).*
  - *There are no documents currently listed under your documents heading on the front page of your site. Where is the literature that makes this product verifiable?*
  - *Below you have a query about uncertainty. I have no idea how you have calculated your estimate of uncertainty. I don't know whether this is a spatial or temporal estimate of the variability or a real effort to tell me how accurate the estimate is.*
  - *Educational materials*
- Combining the data with other datasets:
  - *I would like to be able to combine with other data sets (like SSH or SST or wind or precipitation) on the same time and space scales. How can this be done with OC CCI data products?*

10. Appendix IV: Phase 2, Analysis of Phase 1 Deliverables and Initial Inputs from the User Community on Version 1.0 of the Dataset

Contributors: Samantha Lavender and Shubha Sathyendranath with inputs from various sources.

At the start of Phase 2 the aim had been to cross-check and validate the user requirements that resulted in version 1.0 of the dataset against the feedback being received; both internally and externally. Therefore, information had been gathered through presentations, reviews of deliverables, meetings and emails sent to the initiative help email address.

10.1. Phase 1 Final Annual Review / Phase 2 Kick-Off meeting including Phase 1 deliverables (13/14 February 2014)

Vanda Brotas (FCUL) presented a comparison of the OC-CCI to other datasets containing single and combined missions for a common period. The results highlighted a trend when CCI was compared to the precursor ESA GlobColour product, but not MEaSUREs, which needed further investigation.

Stefano Ciavatta (PML / MAREMIP) presented results from the PML modelling group who were using both the OC-CCI and GlobColour data, including the uncertainties directly. As GlobColour reports larger product errors this means the dataset had a lower assimilation impact compared to the CCI data.

Mike Grant (PML) presentation on known issues:

- *Holes in input data.*
- *Uncorrected bias when new sensors come on-board.*
- *Water classes don't sum to 1.*
- *No uncertainty for atot (as it is a sum of other products) and bbp (due to a lack of sufficient in situ data).*
- *Users can't easily distinguish between different a products.*
- *Uncertainty values exist for NULL data.*
- *Typo in ancillary_variables.*
- *Uses BST rather than UTC in summer – only inside NetCDF [not filename].*
- *Slow download speed, PML currently has a 200 Mb link that is being upgraded to 1 GB/s. Will also transfer to JASMINE/CEMS to provide an alternative download route.*

Comment from Shubha Sathyendranath was that *the NASA applied filtering (e.g. cloud mask) is seen as being too harsh*.

Rosa Barciela (Met Office) presented results from the CMUG activities; further clarified within the Technical note on CMUG ECV Quality Assessment Report:

- *CCI has a higher overall % coverage and lower values; noted that the minimum possible stored value in the GlobColour Chlorophyll-a product is higher than in the OC-CCI dataset.*
- *CCI has a lower % coverage for the SeaWiFS daily data, than GlobColour; because NASA now has a stricter policy for glint masking, which was not applied when the GlobColour GLOB_4KM/RAN dataset was originally generated in 2008.*
- *For January 2003 GlobColour has more pixels in the hook of Africa than CCI*

- *April 2003 – Differences between CCI and GlobColour:*
  - *Arabian Gulf very different between CCI and GlobColour.*
  - *CCI has high patches in the Red Sea (PML is looking at this) and where it joins the Mediterranean (Euphrates)*
  - *CCI high values around New Zealand and stronger Mauritanian upwelling*
  - *GlobColour higher in the Northern North Sea and East coast of Scotland.*
- *Against global fCO2 from SOCAT the RMS error is 71.4 for GlobColour and 68.5 for CCI with 0.36 and 0.43 for the correlation, respectively.*

Recommendations/comments:

- *Remove BST from files*
- *Dimensions and variable names have changed, which causes problems.*
- *Monthly chlorophyll files are not available*

In addition, the GCOS stability requirement was not looked at in Phase 1. This could be addressed by the CMUG who will compare product and model trends.

10.2. Reading Data Assimilation meeting (April 2014)

Although the meeting was technical and directed mostly to atmospheric/meteorological modellers, Stefano Ciavatta provided some overarching bullet points of relevance to OC-CCI:

- *What is the spatial correlation of the stochastic errors of the satellite data?*
- *What is the correlation of stochastic errors between data extracted from different satellite channels?*
- *How to best compute the data errors when satellite data are down-scaled to the (typically lower) resolution of the model grid? This rescaling is currently done, typically, by techniques of "thinning" or "hobbing" the data.*

The mismatch between the resolution of the model and satellite products, the so-called "representativeness" error, that is as important to know as the 'instrumental' error in data assimilation (Error is also called "representativity" or "representation" error: there was no agreement on the proper choice between the names currently in use).

10.3. CMUG feedback within their Phase 1 deliverables (released between August 2013 and May 2014)

D2.2 Technical note on CCI system requirements (v0.7 August 2013)

- Data delivery:
  - It was not clear this will meet the user's needs when users may not want merged products.
  - The dataset would be a good candidate for the Obs4MIPS portal, which was not mentioned.
- Processing approach:
  - Concern that the ECMWF operational fields (not ERA-Interim reanalyses) are used for the atmospheric correction, which was seen as introducing step functions into the products when there are changes in the ECMWF model which affect the total column water vapour.
- Uncertainties:

- The approach could be adopted by other ECVs and it was suggested that this information should also be in an ATBD.
- The uncertainties in areas of sunglint were not addressed and should be
- For the assimilation of chlorophyll concentrations the observation errors used to date are based on a combination of O-B values (Hollingsworth-Lönnberg method) and forecast differences (NMC method). The values provided with the dataset, partly because these don't include representativity error, are not used by assimilation centres. Nevertheless they are a useful input into the estimation of the observations errors.
- Having horizontal error correlations would be of interest for assimilation which was not proposed here.
- The uncertainties provided are also useful for validation.

Technical note on Validation and User Assessment (v0.4 March 2014), focused on the round robin activities:

- Not obvious there was an overall assessment of the final product, which was seen as a major shortcoming.
- Repeat of previous comment about the need to use ERA-Interim.
- Equation (3) in the PVARS-Part1 should be corrected – the second line represents "standard error of relative RMSE (not absolute RMSE)".
- In the atmospheric correction round robin, SeaDAS was showing advantage over POLYMER for SeaWiFS and MODIS processing. Could this be attributed to the use of different auxiliary meteorology and use of different in-situ dataset in the matchup study?
- PVARS-Part 2, Page-14 "comply waters with lw(665)<0.02 (higher scattering waters)" should read lw(665)>0.02.
- PVARS-Part 2, Page-26 "Chi-square test" needs clarification as to why this was defined as a chi-square test in order to avoid the ambiguity. Was the idea to test the Null Hypothesis here?
- While validation of the products was a continuous process, there were still concerns with regards to the under-sampling of the in situ datasets particularly in the low and high productive (chlorophyll-a concentration) regions.
- The units of the chlor_a_rms_uncertainty and chlor_a_bias_uncertainty variables in the OC-CCI NetCDF files were not clear so the documentation should be improved.
- The OC-CCI team's proposition for periodic comparisons of algorithms when there was a significant change to either the in situ observations or retrieval methods, followed by mission re-processing, was commendable. However, there should also be a system to archive the previous version(s) with corresponding training and validation dataset to maintain backward compatibility.

Technical note on CMUG ECV Quality Assessment Report (v2.1 April 2014), initial assessment based on comparison to the GlobColour product, by assimilating both data sets into a coupled physical biogeochemical ocean model [received as part of an oral presentation].

- Product anomalies:
    - Percentage of observations rejected each day had a seasonal cycle, but fairly evenly distributed and so the impact on the final assimilation results was likely to be small.
    - OC-CCI files were a day out during British Summer Time (BST).

- Comparison with GlobColour through assimilation:
    - The biggest difference was the chlorophyll concentration in the sub-tropical gyres, particularly in the Pacific, with lower concentrations in the OC-CCI data, and therefore the CCI-2003 model run. A contributing factor was that the GlobColour products have a higher minimum value, but even taking this into account, the sub-tropical gyres were much clearer in the OC-CCI data.
    - Neither the OC-CCI nor the GlobColour data show a clear global trend over the period, however in the latter years the GlobColour products show a much larger amount of variability, which was believed to be largely spurious due to changes in algorithms and sensor characteristics.
    - At this stage it can be concluded that both data sets were suitable for data assimilation, but neither had yet been definitively shown to be superior to the other, and indeed the different results achieved with each were of scientific interest.
- Formatting of the products:
    - Sinusoidal gridded chlorophyll NetCDF files: Overall, these were convenient to use and fit for purpose. A value for every point on the grid was explicitly stored in the OC-CCI files, whereas with the GlobColour files this information needs to be reconstructed by the user. This makes the OC-CCI files easier to work with and understand, although the trade-off was that data processing takes longer.
    - File format was slightly altered in advance of v1.0 – this was okay when working with test releases, but for a regular release cycle a fixed format would be required.
    - It would also be desirable for single-sensor along-track products to be available alongside the merged daily-average products. The extra time information in these files, as well as the separate error classification for each sensor, would be of potential benefit for data assimilation.
    - Another requirement was for case II waters products (or ideally a single product designed for use in both case I and case II waters), which could be used for validation of, and assimilation into, the Met Office shelf seas model.

Note on Requirements on CCI datasets for inclusion in Obs4MIPs (v0.3 May 2014):

- Data Gridding:
    - Transformed on to a 1 degree grid square through averaging and/or interpolation and then averaged over 1 month; a gridding on a higher resolution (e.g. 0.25 degree) was in general possible and should be considered. It’s also likely that also dataset with higher temporal resolution (e.g. daily) will be desired.
    - For each grid square the fields should be complete (i.e. no data voids) and consideration should be given to ensuring the variables were still conserved in the regridding. It was assumed only observational data (i.e. no model analyses) were included.
    - The associated uncertainties need to be provided on the same grid and care had to be taken to derive these from the level 2 single field of view observations.
    - *CMUG comment: There are several issues to bear in mind when regridding data from level 2 to level 3 especially when dealing with uncertainties. These include maintaining consistency between variables after the regridding, consistently dealing with coastal areas correctly and how to fill data voids. It would be a good idea for CCI teams to share experiences on their*

regridding methodology to avoid and/or document biases resulting from temporal and spatial aggregation.

- Data Format:
  - NetCDF version 3 and must follow the standard NetCDF Climate and
  - Forecast (CF) Metadata convention; in the future it was likely that NetCDF4 format will become the standard for obs4MIP.
  - Must pass a CF compliance check: http://puma.nerc.ac.uk/cgi-bin/cf-checker.pl
  - Each output file must contain a time series of ONLY ONE physical variable. If the entire time series can be stored in less than 2GB, it must be stored in a SINGLE file. If it requires more than 2GB, it should be split into the minimum number of files required, with the size of each file being less than 2GB.
  - Each file should contain a contiguous time series of complete data grid blocks plus all of the required metadata.
  - Each physical variable and coordinate variable must use the specified output/coordinate variable name given in the CMIP5 Requested Output list or should be easily comparable to some derived values from these variables.
  - *CMUG comment: Sharing of experience within the CCI on writing the compliant format datasets from the CCI data would be worthwhile. Feedback on any problems should be given to the Obs4MIPs team.*
- Documentation
  - A short technical note (5 pages max) must be provided with the dataset that conforms to the obs4MIPs technical note template. It should be written bearing in mind the reader will not be familiar with satellite datasets and should make it clear what are the advantages of the CCI datasets with respect to previous datasets already available through Obs4MIPs.
  - *CMUG comment: These technical notes are valuable in their own right to promote the CCI datasets and should be part of the phase 2 deliverables. CCI teams should make them available on their web sites.*
- Representation of uncertainties
  - In the obs4MIP datasets, the 'product uncertainties' was just represented by the variance and number of samples for each grid cells. For CMIP6, the usage of observational ensembles was discussed as an easy way to represent uncertainty information for the user.

10.4. Feedback on version 1.0 of the OC-CCI dataset from users

User feedback (comments/questions from David Ford covered by the CMUG reporting):

- Jean-Michel Andre (Mediterranean Institute of Oceanography, Aix-Marseille Université): needed further information on band shifting and merging. Then, had problems reading the NetCDF files – using a library without NetCDF 4.0 support.
- Francesca Cecinati (ESA / ESRIN): interested in monthly averages at 0.25 degrees resolution.
- Philipp Groetsch (WaterInsight): possibility to download a spatial subset of the OC-CCI data set, e.g. through a WMS or THREDDS server
- Taka Hirata (Hokkaido University): Availability of a land mask
- Dimitra Konsta (Research Centre for Atmospheric Physics and Climatology Academy of Athens): had problems reading the NetCDF files – using a library without NetCDF 4.0 support.

- Mark Payne (National Institute of Aquatic Resources, Technical University of Denmark): Enquired about sub-setting of data – pointed at the OPeNDAP facility.

An explanation of how to check whether support for NetCDF 4.0 is available is now available on the website Frequently Asked Questions (FAQ) page: http://www.esa-oceancolour-cci.org/?q=faq

11. Appendix V: Phase 2, Requirements gathered through the CLEO WORKSHOP

Contributors: Shubha Sathyendranath, Stephanie Dutkiewicz, Cecile Rousseau, Keike Vogt, Nicolas Gruber and Samantha Lavender.

11.1. Why Interactions between the Ocean Colour, Modelling and In Situ Communities are Important.

The ocean plays a considerable role in the global carbon cycle. Globally, oceans act as a sink of carbon dioxide, taking up approximately 30% of the carbon dioxide emitted. It is crucial we understand the role of oceans and the effects that climate change may have on oceans and the carbon cycle. Various methods can be used to evaluate the pools of carbon in the oceans. From in situ sampling to satellite and numerical models, each of these approaches provide a different piece of information in our understanding of the dynamic and variability of carbon pools and fluxes. While in situ data are often used to provide ground-truth datasets used for calibration and validation of satellite data and parametrization of numerical models, the collection of in situ data is labour intensive and can be expensive. Satellite data can provide global datasets but the variables that can be derived from ocean colour are limited to only some of the variables involved in the carbon cycle. Furthermore, ocean colour data can be limited in time and space because of high solar zenith angle, interorbital gaps and the presence of aerosols and clouds that can be considerable in some regions. Finally, numerical models represent a best approximation of our understanding of processes driving the carbon cycle in the oceans but allow for a global representation of variables for which processes are understood well enough that they can be represented in models. The use of satellite ocean colour assimilation can sometimes improve the models and provide an integrated framework in which satellite data are assimilated in models. Numerical models can in turn support the planning and design of field campaigns by conducting simulations experiments of various sampling strategies as well as provide a forecast of the conditions that can be encountered during the sampling campaign.

It was therefore crucial that in the future:

- Quality In situ are maintained for the calibration and validation of satellite data as well as parametrization of numerical models.
- Satellite ocean colour missions are maintained to avoid any gaps in the satellite ocean record.
- Satellite and in situ data are integrated with the use of numerical models to improve our understanding of the carbon cycle and its potential future changes.
- Consistent uncertainties quantification of satellite ocean colour products are provided so that modelling communities can take these into account when using those products for assimilation or validation.
- Interactions between the satellite ocean colour and modelling community continue to strengthen.

11.2. Ocean Colour and Climate

Ocean Colour satellite products offer unprecedented datasets for global and long term monitoring of the oceans. As such these data can be used to help us further understand the ocean and ocean biota role in the earth climate, and a means to monitor changes. As a roadmap to continuing and improving these products the following are important issues:

- There needs to be a concerted, agency level funded, long term system to validate satellite products in-situ. Such a program should include validating all products that will be provided to the user community (e.g. reflectance, Chl-a, IOP) and an in-situ program of validation should also be in place before providing newer products to the users (e.g. functional types). Users were often not ocean colour experts, and require confidence in the products before they will use them. As such it was up to the provider to adequately vet and continuously checking the validity of the products.

- Specifically for the purpose of monitoring the ocean over decades, there needs to be a well-funded, continuous project to provide merged products (i.e. products that encompass different satellite missions). Without such a project, the ability of non-ocean colour specialists to monitor the ocean will be compromised. Such a project should be constantly updating and re-evaluating the products, especially as new satellite missions (e.g. Sentinel 3) come online.
- Numerical biogeochemical and ecosystem models were one of the "users" of ocean colour products. Such models were often interested in the carbon cycle, and usually represent phytoplankton and zooplankton in terms of their carbon content. This leads to a disconnect between what ocean colour products provide (e.g. the most commonly used product was Chl-a) and what modellers specifically "need". Though there was a strong effort from the ocean colour community to take the step towards providing carbon (C) products, there needs to be some level of caution. For instance, most ecosystem models now include dynamic Chl:C for the modelled phytoplankton. There was likely less error in these numerical model Chl:C ratios than there are in Ocean Colour product Chl to C (or Rrs to C) conversions.
- A strong dialogue between ocean colour community and modellers would facilitate an agreed-on standard for use of ocean colour products to avoid some of the issues mentioned above. Tutorials to target modelling audiences, workshops to agree on standards, documentation that do not reply on ocean colour jargon are some of the possible ways to facilitate better co-operation between the two groups.
- On the other hand, numerical models are now beginning to include output that relate more directly to ocean colour products (e.g. Rrs – Dutkiewicz et al., 2015; Baird et al., 2016; IOPs – Fujji et al 2007; Gregg and Casey 2010; Dutkiewicz et al. 2015; Ciavatta et al. 2016). Given that the models do not have the issues of missing data (clouds, satellite repeat cycles etc.), and have 3 dimensional information, they could be useful laboratories to explore uncertainties in ocean colour products, help in the development of new algorithms etc. Even without these directly links to ocean colour products, models can and have been used to provide information to the ocean colour community (e.g. Henson et al. 2010 provided estimates of the length of time before definite anthropogenic driven trends in Chl-a could be detected). The continued and improve use of models in helping address ocean colour issues will require communication between modellers and ocean colour community. Targeted workshops and funding opportunities for such collaboration was urgently needed to advance this promising direction.

11.3. Next Generation Ocean Colour Observations for Climate Studies

Recommendations:

1. Towards a 4-dimensional ocean colour field for climate modelling

Climate models are highly reliant on observations in order to constrain the processes represented and the patterns of variability produced. In recent years, the ocean community has made a tremendous effort to better understand the ocean using in situ observations made by autonomous platforms such as floats and gliders. On September 1st, 2016 the Argo initiative had 3739 floats operating in the global ocean and providing information on temperature, salinity and ocean currents, most of which are also measurable for the surface ocean from space. With an increasing number of Biogeochemical-Argo floats being released (Biogeochemical Argo Task Team, 2016), additional measurements of chlorophyll, backscattering, CDOM and irradiance will soon be available at a high frequency and with good global coverage. It seems timely ***to calibrate and improve all existing satellite products with the Argo data***, and to link into existing and currently planned activities in order to study ***the link between surface ocean conditions and the depth profiles of these variables***. We recommend for ESA ***to invest in cross-disciplinary projects*** in order to ***develop statistical algorithms that furnish users with preliminary 4-dimensional products of ocean observables***. The products we envisage should be related to those specified in the Global Ocean Observatory System list of ***Essential Ocean Observables*** to first priority, and products related to ***ocean productivity and export, as well as plankton biogeography***.

2. A European EXPORTS initiative?

Recently, a wide range of products capturing aspects of the biological carbon pump from remote sensing observations has been made available for the modeling community. The range of products comprises estimates of **net primary production** (e.g. Behrenfeld and Falkowski, 1997; Behrenfeld et al. 2005, 2013; Westberry et al. 2008), ***phytoplankton carbon biomass*** (e.g. Behrenfeld et al. 2005; Behrenfeld and Boss, 2006; Siegel et al. 2005, 2013; Dall'Olmo and Westberry, 2009; Brewin et al. 2012), ***phytoplankton functional types*** (e.g. Alvain et al. 2005, 2008; Nair et al. 2008; Bracher et al. 2009; Hirata et al. 2011), ***phytoplankton loss rates*** (e.g. Behrenfeld et al. 2010, 2013; Siegel et al. 2014), ***suspended particulate organic carbon*** (e.g. Stramski et al. 1999; Mishonov et al. 2003), ***particle size spectrum*** (e.g. Ciotti and Bricaud, 2006; Kostadinov et la. 2009, 2010; Mouw and Yoder, 2010; Bricaud et al. 2012; Roy et al. 2013), and ***rates of export fluxes*** (e.g. Laws et al. 2000; Henson et al. 2011; Siegel et al. 2014). Thus, ***a full quantification of the upper ocean biological carbon pump using remote sensing products*** seems no longer out of reach and would be essential to better constrain marine ecosystem and climate models.

A recent initiative of the oceanographic community in collaboration with NASA is specifically targeted at the investigation of the biological pump (**NASA-EXPORTS**, Siegel et al. 2016). This initiative combines in situ and remote sensing observations with modeling studies in order to better understand and constrain the biological carbon pump in the ocean. Since the initiative has been conceived jointly by a large network of modelers, remote sensing experts and observational oceanographers, it is specifically tailored at improving our quantitative understanding of all aspects of the biological pump from carbon uptake and the conversion of inorganic to organic nutrients by marine autotrophs to the cycling within the marine food web and to the degradation of organic matter into its inorganic constituents by bacteria, with a particular emphasis on plankton biogeography and diversity, and the life style of zooplankton (diel vertical migration). This initiative will consist of targeted cruises in the North Atlantic and North Pacific and the combination with remote sensing and modeling experiments targeted at a mechanistic understanding of the measurements and the improvement of satellite algorithms.

In our view, ***a European participation in EXPORTS***, either through an ***extension of planned activities to other under-sampled ocean basins*** (South Pacific, Southern Ocean, Indian Ocean), or through ***close cooperation*** in terms of missions and sensors inter-calibration for those sensors covering the target areas at the time of the activities would be exactly what climate modelers would need to better constrain the biological aspects of the ocean carbon cycle for the present and future ocean.

3. A European contribution to 'Who is where and why?'

In recent years, reports of changes in marine ecosystem structure due to climate change have been increasing rapidly, with changes comprising species and regime shifts, the apparition of invasive species and the expansion of entire biomes (e.g. Reid et al. 2007; Beaugrand et al. 2009; Rivero-Calle et al. 2016; Barton et al. 2016; Poloszanska et al. 2016 and references therein). The ***drivers of these changes in species composition and diversity*** remain poorly understood, and the ***consequences for marine ecosystem function and the provision of essential ecosystem services*** are largely unknown. Remote sensing observations are the only source of information that would facilitate ***consistent long-term monitoring of remote ocean regions*** that are not regularly frequented by commercial and scientific vessels. Hence, in our view, ESA and other space agencies have a societal responsibility when it comes to the monitoring of marine diversity on planet Earth. In fact, we would consider it of paramount importance for ESA to make ***the quantification of marine ecosystem structure*** and ***the definition of marine ecoregions*** and ***hotspots of ecosystem change*** a main focus for their future research and development.

While it will not be possible to monitor the dynamics of individual species from satellite, remote sensing information may provide us with information on the state of the marine ecosystem through the observations of plankton functional types, i.e., groups of species with a similar ecological function and a significant contribution to global marine biomass on the regional to global scale (Le Quéré et al. 2005). Observations of PFTs or their specific pigments paired with those on environmental conditions may allow for the ***development of empirical models of community composition from remote sensing*** that may be useful to monitor ***climate change impacts*** and ***geographical shifts of different ecoregions***. As the list of references in appendix 2 demonstrates impressively, Europe is at the forefront of research on the detection of phytoplankton functional types from space and the impact of climate changes on marine diversity and biogeography. However, ***the many different satellite products differ substantially in their patterns***, and the groups of organisms represented in different PFT products vary between the products.

In our view, ESA could be an essential partner for ***inter-calibration of different satellite products***, with the view of ***providing new standard satellite products on marine diversity and ecosystem structure*** in the longer term. To this end, ESA should invest in ***hyperspectral instruments*** and actively foster the inter-calibration with traditional PFT estimates from multi-spectral methods. Like, chlorophyll, ***marine ecosystem composition from space should be one of the focus areas of future ESA research***, and ***a new suite of high quality PFT products*** should be developed. To this end, ESA should ***combine environmental data on SST, PAR, CDOM, POC, as well as novel estimates of variables part of the carbonate system (e.g. ALK) known to affect phytoplankton physiology with PFT detection algorithms and new statistical techniques***, in order to better constrain phytoplankton biogeography, its drivers and the ecological niches of key organisms. This is to inform society on the magnitude of change in biodiversity and distribution patterns, but also to inform existing ocean and climate models of marine biogeography and to better constrain the biological pump. But even more, this is ***to inform also the development of a new generation of global marine ecosystem models based on functional traits and with a better resolution of plankton diversity***, i.e., models that are targeted at better constraining the impact of climate change on ecosystem service provision and higher trophic levels and fisheries. These novel models have not yet been included into global Earth system models and the IPCC assessment due to computational constraints on their use in climate projections, but are likely to be used as such within the next 5-10 years.

To this end, it would be essential to develop PFT algorithms by combining remote sensing variables with the existing data on phytoplankton biogeography available from projects such as ***MAREDAT*** (Buitenhuis et al. 2013), the past and ongoing ***TARA-OCEANS*** expedition (de Vargas et al. 2015; Guidi et al. 2016), or the ***AMT transects*** (Robinson et al. 2006), which all monitor different aspects of phytoplankton diversity and biomass distribution. These data are ready for inter-comparison with PFT estimates from space for the development of better and comparable satellite products. In order to understand the drivers of marine community composition, information that links community structure with environmental conditions is essential. This highlights the importance of continued improvement of existing and currently developed satellite products related ***to (surface) ocean temperature, irradiance, nutrients, turbulence (eddies, mesoscale variability), carbon pools (POC, CDOM, DOC) and carbonate chemistry variables (Alkalinity, DIC, pH, ...) and their variability on time scales ranging from hours to decades***. These variables are known to play an important role in the metabolism of autotrophs responsible for ocean productivity, and should be further improved and carefully validated with in situ observations.

5. Long- term consistency in data records

Last but not least, climate models attempt to quantify global biogeochemical cycles in general and the carbon cycle on time scales of ***decades to millennia***. Thus, ***long-time time series of observations*** on these time scales are essential for model development and validation. Thus, future satellite products need to be ***consistent*** with existing and past products, and the ***continuity of existing (satellite) time-series*** must be guaranteed. ESA should take a lead role in providing a ***long-term, continuous record of ocean colour and derived products with consistent instrumentation***.

11.4. End User Requirements

The end user requirements from the CLEO workshop were summarised in the discussion session as being:

- From the (extended) climate community:
    - Impact of climate change on marine ecosystems (Intergovernmental Panel on Climate Change, WG1 and WG2). Essential for the vast areas of the open ocean and the only way to measure diversity at the larger scales.
    - Assessment of climate conditions in Earth System models and process-based model parameterizations.
    - “Boundary” conditions in high-resolution, physics-only models (radiative heat flux propagation, marine aerosol).
    - Estimation of heat and carbon budgets, organic aerosol, air-sea/land-ocean fluxes and export production (also through the use of reanalyses).
    - Two-way interaction: models can inform the EO missions on required spatial and temporal resolution and length of the time series to capture climatic trends and variability.
    - Common/cross-cutting needs: Many requirements of marine environmental services were the same as climate requirements. Exchange/transfer between environment and climate services; recommend agency-level consistency in processing chain for environment and climate services.
- Continuity and consolidation:
    - The curation of ocean colour data was critical for the long-term detection of climate impacts.
    - OC-CCI initiative if not continued, would constitute a waste of the initial investment and the achieved consensus on fundamental principles:
        - Time series should be protected from gaps in data stream (key to have two sensors in orbit at the same time).
        - Modifications required in response to reprocessing, new sensors and advances in science.
        - The production of a great service came with responsibility: unwise not to spend a small fraction of mission budget on high-quality products and their access (user friendly portals, user support, tool boxes, avoid misuse of data).
    - The continuous and sustained assessment of the quality of the products was the only guarantee that the small trends that have been observed at global and regional scales are significant and should be reproduced in the historical simulations.
- Validity and accuracy:
    - The level of accuracy **MUST** be high to detect the smallest trends.
    - Too few data on certain products to assess their validity (PFT distributions).
    - Merged ocean colour products require a concurrent effort to maintain a selected set of long-term oceanic time series for validation:
        - Synergy with observations: bring together in situ observations, satellite data and models to study ocean processes relevant in the climate context.
        - Add in new autonomous and remotely-operated vehicles (gliders and bio-Argo)
        - In situ validation and calibration programmes must keep pace with enhanced product line.
- New products:
    - Development of targeted regional indicators: case studies to illustrate the value of ocean colour as ECV (small island states, Arctic).
    - Generation of sub-surface fields with the adequate uncertainty estimation (would contribute to carbon export and fluxes and the use of prescribed fields in physics-only models).
    - Improved algorithms (cross-cutting applications):
        - key areas;

        - species/types important in the climate context;
        - ocean economy (fisheries, aquaculture).
    - Exploit LIDAR data for Ocean Colour products in marginal ice zones (emphasis on the type of measurements and not the specific.
- Inter-operability and societal needs:
    - Common messages are important: GCOS / Global Ocean Observing System (GOOS) to work together with respect to commonalities in ECVs and Essential Ocean Variables requirements.
    - Ocean colour was also recognised as an Essential Ocean Variable, which admits a broad range of applications, beyond climate. The GOOS biogeochemistry panel takes responsibility for Ocean Colour, but GOOS recognizes the cross-cutting nature of Ocean Colour.
    - Development of cross-sectorial linkages:
        - ocean colour and marine ecosystems;
        - socio-economic impacts;
        - fisheries implications;
        - transfer from science to evidence base for policy makers.

12. Appendix VI: Phase 2, Summary for the product specification

- Dataset anomalies:
    - All observations need to have an uncertainty for them to be assimilated. [Was addressed in Phase 2, and resolved for v2.0]
    - OC-CCI files were a day out during BST. [Was updated in the code, didn't occur during the v2.0 production]
- Data delivery / documentation:
    - Improvement in the explanation of the units for the chlor_a_rms_uncertainty and chlor_a_bias_uncertainty variables. [Available on the website FAQ page]
    - Release through Obs4MIPs? In particular, produce a short technical note in the short term. [In progress, linked to v3.0]
- Processing approach:
    - Use of ERA-Interim rather than operational ECMWF data for product generation. [Discussed in detail at the Phase 2 KO, reanalysis dataset will be used; ERA-Interim or NCEP]
    - Requirement from the Met Office / CMUG for single sensor products. [In production for MERIS]
    - Requirement for case II waters products (ideally a single product designed for use in both case I and case II waters). [Addressed in Phase 2, within version 3-0]
- Round Robin (RR) activities:
    - Overall assessment of the final product? [Was performed within the error characterisation and was continually being worked on for each new release; in the 2nd RR activity the *in water* RR used the outputs of processing (after the AC RR has chosen the best configuration) rather than the *in situ* database]
    - Under-sampling of the in situ datasets, particularly in the low and highly productive regions? [Improved for highly productive waters within the *in situ* database v2.0, but will still be under-sampled in low productivity regions – overall, there was a lack of data community collected data in the open oceans]

- System to archive the previous version(s) with corresponding training and validation dataset to maintain backward compatibility? [Being worked towards for AC and already exists for *in situ* RR]

- Uncertainties:
  - Uncertainties in areas of sunglint – Are sunglint contaminated pixels used in the calculation of the uncertainties and can the data be separated to show (not show) higher uncertainty resulting from the presence of sunglint? [Within Müller et al. (2015) and the Phase 1 Product Validation and Algorithm Selection Report (PVASR)]
  - Can we understand more than uncertainties? Types of error mentioned by the modelling community as important include the spatial correlation of the stochastic errors, correlation of stochastic errors between data extracted from different satellite channels and the representativeness or representativity error. [Discussed at the Informal Workshop of the Climate Research Group, will be discussed with Chris Merchant (SST CCI) and Nick Rayner (UK Met Office)]