Climate Office
Harwell Centre

Atlas Building,
Harwell Science & Innovation Park
Didcot, Oxfordshire, OX11 0QX, UK
T +44 (0)1235 567900
climate.office@esa.int

DOCUMENT

CCI Second Collocation Report

Prepared by	**Pascal Lecomte and Cat Downy**
Reference	**CCI-PRGM-EOPS-TN-11-0029**
Issue	**2**
Revision	**0**
Date of Issue	**February 7th, 2012**
Status	**N/A**
Document Type	**RP**
Distribution	**Public**

European Space Agency
Agence spatiale européenne

APPROVAL

Title CCI Second Collocation Report	
Issue 1	**Revision** 0
Author	**Date** October 19th, 2011
Approved by	**Date**

CHANGE LOG

Reason for change	Issue	Revision	Date

CHANGE RECORD

Issue 1	Revision 0		
Reason for change	**Date**	**Pages**	**Paragraph(s)**

Table of contents:

1 SCOPE

This document summarises the outcomes of the Second Collocation meeting of the ESA Climate Change Initiative, held at ESA ESRIN, Frascati, Italy on 12th to 14th October 2011.

The collocation meeting brought together representatives of all eleven CCI project teams initiated in 2010 as well as the three new projects that are starting in autumn 2011, to discuss issue of common interest to all. The collocation had two major objectives:

- To take stock of progress during the first year of CCI projects
- To start preparing for the next phase of CCI programme (3 years)

The output of the collocation is recorded in this report as a collective assessment of the status of the CCI and a series of key actions, as formulated by the collocation participants, drawing upon the collective expertise of all CCI projects teams.

This report is intended to assist the CCI teams to pursue their projects and generate ECV data products in a consistent manner, as explicitly required by GCOS, and to prepare for phase 2 of the CCI programme.

2 INTRODUCTION

The CCI programme objectives and scope are described in the document, "ESA Climate Change Initiative: Description [EOP-SEP/TN/0030-09/SP]".

The work to be carried out on each ECV is specified in the Statement of Work for the CCI, "ESA Climate Change Initiative Phase 1: Scientific User Consultation and Detailed Specification [EOP-SEP/SOW/0031-09/SP]".

The project scope and team composition for each of these projects is described in the document "ESA CCI Projects Description" (all documents can be found on the CCI website).

The first tender of the CCI programme resulted in projects for the following ten ECVs:

GCOS ECV	CCI Project	Science Leader
A.6	Cloud_cci	Deutscher Wetterdienst (*R.Hollmann)*
A.9	Ozone_cci	BIRA-IASB (*M. Van Roozendael*)
A.10	Aerosol_cci	DLR / FMI (*T Holzer-Popp / G.De Leeuw*)
A.8	GHG_cci	U.Bremen IUP (*M.Buchwitz)*
O.3	Sea_Level_cci	LEGOS-CNES (*A Cazenave*)
O.1	SST_cci	U. Edinburgh (*C Merchant*)
O.6.1/2	Ocean_Colour_cci	Plymouth Marine Laboratory (*S. Sathyendranath)*
T.3.1/2	Glaciers_cci	U. Zurich (*F.Paul)*
T.6.1	Landcover_cci	Université Catholique de Louvain (*P.Defourney*)
T.10	Fire_cci	U. Alcala (*E.Chuvieco)*

At the same time, through a separate ITT, the cross-cutting project, 'Climate Modelling Users Group' was set up with the aim of ensuring integrated feedback between the ECV projects and the global climate modelling community. The CMUG group provide a structured forum for reviewing key ECV documents, discussing intercomparison methods, error characterisation, etc and guiding the projects to achieve the set performance targets for each ECV.

In 2011 three new projects were set up and although none had officially kicked off by the time of the second collocation meeting it was still very beneficial for them to attend the meeting. The following ECVs are being addressed by the new projects:

GCOS ECV	CCI Project	Science Leader
O.5	Sea_Ice_cci	NERSC, *(S. Sandven)*
T.4	Ice_Sheets_cci	DTU Space *(R. Forsberg)*
T.11	Soil_Moisture_cci	TU Vienna *(W. Wagner)*

Each CCI project team typically includes experts from ten or more research organizations, including a Science Leader and a Project Manager, and contains three sub-groups. One has specialist scientific expertise in EO, another is specialised in climate research and modelling, and the last group contains system engineering experts.

Each science leader will ensure the overall scientific integrity of the project throughout the next three years. The science leader will also ensure that each CCI project maintains effective working links to the appropriate international climate science programmes, initiatives and projects, and to other CCI project teams. Each science leader is directly supported by a project manager who will ensure communication within the project team, maintenance of schedule, tracking of actions, deliverables and reporting to ESA.

The projects are still in phase 1 of the programme but have achieved many of the key deliverables set out in the first Statement of Work. The CCI project deliverables (data, quality, cal/val, documentation, review, open access) have been specified in accordance with the "Guideline for the Generation of Satellite-based Datasets and Products meeting GCOS Requirements" (GCOS-129, March 2009). All completed documents can be found on their websites.

Attendees of the first collocation meeting put together some Project Guidelines with specific recommendations for phase 1, 'ESA CCI Project Guidelines' [EOP-DTEX-EOPS-SW-10-0002]. These have guided the first year of the 10 initial ECVs and will hopefully inform the development of the 3 newer projects.

The next phase for the CCI and the projects is to complete phase 1 and prepare for phase 2. The second collocation meeting dedicated significant time to discussing system requirements and system engineering, both in plenary and in drafting groups. It is intended that the recommendations from both collocation meetings will inform the Statement of Work for phase 2.

3 SPLINTER GROUP REPORTS

3.1 CCI Science Agenda

The contributors for this chapter are:
Gerrit de Leeuw (Aerosol_cci)
Rainer Hollman (Clouds_cci)
Roger Saunders (CMUG)
Emilio Chuvieco (Fire_cci)
Frank Paul (Glaciers_cci)
Michael Buchwitz (GHG_cci)
Sophie Bontemps (Land Cover_cci)
Shubha Sathyendranath (Ocean colour_cci)
Michel van Roozendael (Ozone_cci)
René Forsberg (Ice sheets_cci)
Anny Cazenave (Sea_level_cci)
Chris Merchant (SST_cci)
Wolfgang Wagner (Soil moisture_cci)
Pierre-Philippe Matthieu (ESA)
Carolin Richter (GCOS)
Jean-Louis Fellous (COSPAR)
Mark Doherty (ESA)
Cat Downy (ESA)

3.1.1 How can the project teams maximise the science impact of the CCI?

A number of suggestions were put forward to increase the science impact of the CCI programme. The first set of bullet points describes actions that could be undertaken by the science leaders:

- The projects have already agreed, as part of their contracts, to publish their results in peer-reviewed journals that are cited by the IPCC assessment reports. This will provide the biggest impact in terms of advancing scientific understanding of the climate system and supporting the UNFCCC. It is recognised that the major impact of the CCI will only be seen in AR6, however some projects do have papers in press in time for AR5.
- A community paper in the Bulletin of the American Meteorological Society to describe the programme. **Action: Roger Saunders and Cat Downy to coordinate inputs from the science leaders.**
- A special issue of a journal to jointly publish the projects' Round Robin results. Each paper could have a common section in which the RR exercise is described in a comparable way between the projects. Carolin Richter expressed interest in being involved and it was agreed that an introductory paper from a GCOS perspective would be useful to add context. Suggested journals: Remote Sensing of the Environment, COSPAR Space Research.
- A session at a conference (e.g. EGU, AGU, IGARSS), preferably in Europe to get a critical mass of science leaders to attend but other conferences to be considered. **Action: Wolfgang Wagner to explore the option of having a session or union meeting at EGU 2013.**
- The CCI data sets are to be put on a central archive and published alongside the CMIP5 (Coupled Model Intercomparison Project, Phase 5) and related international climate data archives (e.g. through ESGF, the Earth System Grid Federation).

> R1: A community paper is to be written by the science leaders; a joint CCI session should be organised at an international conference; the CCI data sets are to be made available alongside international climate data archives (e.g. CMIP5/ESGF); a special issue of a journal publishing CCI Round Robin results is to be investigated.

3.1.2 How can ESA maximise the science impact of the CCI?

- A CCI science brochure, to be about 30 pages long, with 1-2 pages detailing each project to be produced in time for the ministerial.
- CCI iPad app, initially to help communicate with politicians and decision makers at high-level events, such as the COPs and at conferences. It could then be developed as a publically accessible app.
- Report back from the projects to other sections of ESA, to give guidance to EO missions on gap analyses and recommendations for future missions (e.g. the data used in the Fire CCI comes from sensors that were not designed for that purpose).
- There are possible links with WCRP research strategy that could be strengthened but this still ambiguous. A**ction: Carolin Richter to look into what links can be strengthened.**

R2: ESA will coordinate the production of a CCI brochure and an iPad app, with input from the projects.

3.1.3 How can we exploit the interdisciplinary nature of the CCI programme?

The science leaders wanted to make it clear that the maximum science impact lies with exploitation of climate data sets, not with generating the data sets, and at present there is little opportunity within the CCI to address science issues since ESA's focus is on data generation (with the exception of CMUG). There was agreement amongst the science leaders that addressing interdisciplinary science questions and producing consistent ECVs was one of the key benefits to the CCI.

Avenues to explore:

- Once the projects have a better idea of what their products will be it would be useful to develop a more detailed science agenda with key scientific questions. It was felt there was most to gain from studying processes and feedbacks in the climate system that require multiple CCI products (e.g. Cloud_cci and Aerosol_cci to quantify radiative forcing).
- This detailed science agenda could then be used to apply for funding from other sources, e.g. within ESA such as the Support To Science Element and Data User Element, or from the EU or EC.
- The possibility for a 'rapid response fund', i.e. flexible funds to respond to relevant events and emerging issues, e.g. the ESA-NASA working group on ice mass balance.

The group also looked at what could be done to improve interdisciplinary fertilisation during CCI meetings and suggested the following:

- Allow more time for science discussions among teams.
- Allow time for brainstorming of shared science issues.

It was proposed that the next CMUG Integration Meeting should have much more of a focus on the science; science leaders will be more involved in setting the agenda with CMUG. Location: Toulouse, 14-16 May 2012.

R3: Expand the CCI Science Agenda developed at the last collocation to develop key scientific questions that exploit the data products and involve multiple ECVs.

3.1.4 *How can we make the best use of the time allocated for science in Phase 2?*

As noted above, a key attribute of the CCI is that it brings different ECVs together to work as a 'super team', but the science leaders feel that they haven't been able to exploit this advantage yet. This is due to a number

R4: The next CMUG Integration meeting should have dedicated time for the science leaders to get together to focus on the Science Agenda. They should have more input in setting the agenda for the meeting.

of reasons, namely the lack of resources (money, time, people) for additional work within the CCI, given the heavy commitments under the existing programme. Phase 2 offers an opportunity to revisit how the science leaders operate together and how their time can be best spent on the science, rather than with other aspects of ECV generation.

The science leaders made suggestions on how they think time for science could be best allocated in Phase 2:

- Opportunity for more science in Phase 2 (i.e. less technical deliverables and more science deliverables).
- Budgeted money for the 'rapid response fund', noted above.
- Incorporate climate services concept into climate research group.
- Facilitate smoother programme delivery (e.g., easier access to ESA data).
- Create longer time series with overlapping years between CCI's; need a golden year (suggestion 2008).
- CCI for Sentinels, preparatory activities required?

While taking these points into account and encouraging the enthusiasm the projects have for working together, ESA was clear to point out that the scope of the CCI programme is rather strictly defined. The suggestions could be supported by ECV data products but there are a number of organisations that are better suited to funding some of these ideas, such as FP7, Horizon 2020, national research councils, joint projects initiative, etc.

3.2 Algorithms and Performances

The contributors for this chapter are:
Ronald Van der A (Ozone_cci)
Gilles Larnicol (Sea-Level_cci)
Bob Brewin (Ocean-Colour_cci)
Tobias Bolch (Glaciers_cci)
Claus Zehner (GHG_cci)

3.2.1 Status of round robin for each CCI?

The status of each ECV with respect to progress on the round robin is mixed. All projects are preparing for, or are in the process of, their round-robin and selection inter-comparison, even the ones that have a pre-selected algorithm (performing module testing). Some ECVs appear advanced in their round-robin (particularly the ocean ECVs, e.g. Sea-Level, Ocean Colour and Sea Surface Temperature) and some have yet to start (e.g. Fire and Glaciers) as preparation has taken a lot more time than expected and not all projects have started at the same time.

3.2.2 Was there enough time?

There was a consensus among all ECVs that there was not enough time. All ECVs were in agreement that more time is required to produce a thorough and critical inter-comparison. However, for many ECVs the extra time needed to do the round-robin may not lead to a delay in the overall project schedule. For example, the Sea-Level and Aerosol ECV are taking slightly longer to complete their round-robin, but had plans to recover this time (e.g. reduced time for data processing).

All ECVs are in clear agreement that the round-robin and selection inter-comparison should continue throughout the CCI programme. The CCI strategy has to be open to the possibility that better algorithms will emerge in the future, and the future stages of the CCI should include periodic re-evaluations of algorithms, adoptions of new algorithms and re-processing of data archive, as and when is necessary. There has to be the flexibility to include, in timely manner, new auxiliary data becoming available from new sensors (e.g. Sentinel). This recommendation has clear implications for software engineering and data merging.

R1: The round-robin and selection inter-comparison should continue throughout the CCI programme to include periodic re-evaluations of algorithms, adoptions of new algorithms and re-processing of data archive, as and when is necessary.

3.2.3 Was there a common environment?

There appears to be a common environment developed (or being developed) for each ECV (e.g. all using the same Level 1 and auxiliary data for processing). It is of great importance to maintain traceability in the round-robin planning and execution, particularly with regards to continuing the round-robin and selection inter-comparison in the future stages of the ECV project. The Product Validation Plan (PVP) and Product Validation and round-robin Selection Report (PVASR) will help to document methodologies and established standards.

R2: Maintain traceability in the round-robin planning and execution.

3.2.4 Were there examples of external collaboration?

In general, all ECVs were open to external collaboration. There were some examples of strong external collaboration (e.g. Ocean Colour and Fire), but also some cases where external collaborators were not so willing to get involved (e.g. due to lack of resources). Based on the success of the CCI, there could be more external collaboration in the future, potentially with the addition of new algorithms or a consensus toward a standardized approach to the round-robin and intercomparison.

3.2.5 Selection, evaluation criteria, algorithm scoring and ensemble data-sets

In general, the round-robin planning has been well documented (as requested in Statement of Work, via the PVP and PVASR). Algorithm scoring was dependent on the variable being assessed and on the specific ECV. Some ECVs have a fully automated validation procedure, while others are based on both qualitative and quantitative assessment criteria. Rather than providing an ensemble of datasets, for most ECVs it appears preferable to have one.

3.2.6 Who did the algorithm selection?

For several ECVs the algorithms are being selected internally by group consensus but with external review and feedback (e.g. Ocean Colour). In other ECVs, such as Sea-Level, the algorithms are being chosen externally of the ECV group. Real blind testing was only done in one case of the Sea Surface Temperature ECV as they have access to a huge wealth of data.

3.2.7 What did each consortium gain?

All CCI teams are in agreement that a lot was gained in conducting and planning the round-robin. By doing this exercise different experts have been pooled together, which would not have happened without the CCI. Through the intercomparison, the advantages and disadvantages of specific algorithms and single modules (e.g. variability in performance depending on variable, or temporal and spatial scale) were revealed leading to further insight into the different approaches being used. This allows for further algorithm improvements, refinements, and selection. For some ECVs, further algorithm improvements have been achieved by combining the best features of various algorithms.

Furthermore, standard procedures and round-robin protocols that are being developed within each ECV, together with international approval, can lead the way in international algorithm development.

3.3 Data Delivery, Quality and Feedback

The contributors for this chapter are:
John Swinton (Ocean_colour_cci),
Gary Corlett (SST_cci),
Martin Bachmann (Fire_cci),
Carsten Brockmann (Landcover_cci and Ocean_Colour_cci)
Joerg Schulz (EUMETSAT)
Bojan Bojkov (ESA),

The discussions for this chapter took the form of an open session with three seed presentations, followed by a review session with a reduced number of participants tasked with drafting this document.

3.3.1 Input Data Quality

For climate it is essential to implement good practice for data quality issues throughout the process of ECV generation. One of the key elements of data quality is the traceability of the data including its pre-launch (laboratory) calibration, in-flight calibration (both on-board and vicarious) and field of view, as well as information on the platform, including the basic orbital parameters. Information is also required on the data processing: Level 0 to Level 1b to Level 2; the corrections applied; geolocation; auxiliary files; and possibly the open source code. Much of this guidance is laid out in CEOS' Quality Assurance Framework for Earth Observation (QA4EO).

All this information should be updated, as far as it's possible, prior to each reprocessing, including the sensor characterisation report, which comes under the responsibility of QWG (Quality Working Group) and should be available online. The whole process must be responsive.

There is a need to distinguish between the product version and software version; a versioning system is needed for ESA products and possibly a Digital Object Identifier (DOI) too.

Calibration/validation underpins climate data and requires independent, high quality, well-characterised, long-term, geophysically representative, continuous measurements.

R1: As recommended by CEOS, space agencies should maintain a dedicated website for routine sensor monitoring, data quality reports (plus summary) and uncertainty characterisation. The site should contain complete knowledge of the sensor, platform and data processing so that the information is traceable, as recommended in QA4EO.

R2: All ECV data products shall contain quality indicators as well as information on uncertainties. In addition, all products should have a clearly defined versioning system and should have a unique DOI for traceability.

3.3.2 ECV Data Product Delivery

It was noted that most CCI projects have sourced ESA data outside of the official ESA data archives, as it has proved harder and more time consuming to get hold of than originally envisaged. It is therefore recommended that the CCI programme should ensure that all EO data is easily available online by e.g. FTP/HTTP etc. or via appropriate media.

R3: All ECV data products should be easily available through a common portal that supports (a) multiple download mechanisms (e.g. FTP/HTTP and appropriate media for large volumes), (b) subscription services and (c) product sub-setting.

Information flow from agencies to users needs to be improved, and there is an opportunity for the CCI programme to do this. There should be a dedicated website for routine sensor monitoring, data quality reports (plus summary) and uncertainty characterisation. Product User Guides (PUGs) and simple README files should be available for each product.

Robust, responsive, regular reprocessing is needed. The current processing model operating within the ESA ground segment is not suitable for the CCI. A User Processing Facility (UPF) would facilitate Sentinel processing owing to the large data volume. This approach should 'bring the algorithms to the data'.

R4: ESA should consider establishing a User Processing Facility (UPF) for rapid processing and reprocessing of ECV data products. The system would allow users to login remotely and run their own software on the UPF.

The provision of sub-setting (i.e. MMD approach of SST_CCI) for rapid evaluation and for calibration/validation is recommended.

3.3.3 Feedback

Feedback at all levels of the production chain is essential; GCOS has an explicit requirement for user feedback facilities in its product guidelines. Level 1b should be the starting point for CCI processing for optical/IR/MW domains. There are a number of suggestions by the group for where feedback could be improved for the CCI:

- Regular reporting of data quality issues (ESA missions) is needed between the QWGs and CCI teams.
- The CCI teams need a better and more regular exchange of information.
- Best practice protocols should be used for radiance inter-comparisons.
- It would be useful to know how Level 2 products get updated.
- A mechanism for feedback for 3rd party missions should be defined.
- CCI teams need feedback on CCI data products from users outside of CCI.

R5: A routine mechanism should be established for feedback between CCI project teams and QWGs responsible for ESA and 3rd party mission Level 1b data products.

R6: All CCI project teams need to regularly exchange information on data quality.

3.4 System Engineering

The contributors for this chapter are:
Martin Boetcher (Landcover_cci)
Mike Grant (Ocean Colour_cci)
Tazio Strozzi (Glaciers_cci)
Meinhard Wolfmueller (Ozone_cci)
Óscar Pérez Navarro (Fire_cci)
Günther Lichtenberg (GHG_cci)
Anne Chadwick (ESA)
Jose Ramos Perez (ESA)

3.4.1 Round robin processing

The teams summarised the work they have done to process the round robin data in the discussion group. Many used dedicated computing clusters and no one has use of a supercomputer; most jobs fitted on a single cluster. There have been a number of problems with the round robin processing, namely data transfer issues, problems with handling the bulk of the data, the location and latency coming from providers and the fact that error handling is often done manually. This was because correcting a problem often required certain expertise, e.g. SST had errors in more than 50% of their months. It was noted that the research group behind MyOcean run a comparable operational system and that an agile methodology is useful for development. Michael Buchwitz (GHG_cci) suggested that some ECVs could make use of industrial-style processing but this is not the case for all projects so flexibility is needed in defining operational environments.
Some requirements are not yet reflected in the Round-Robin/Prototype systems such as interface requirements, some operational requirements, etc.

R1: The SEWG should review requirements and analyse requirements across the CCI to derive commonalities from the preliminary system analyses.

A preliminary list for R1 was drawn up, which looked at how different system levels are defined and which levels have common requirements.

- Host infrastructure, hardware, data centre, and private clouds.
- Data storage & access
- Processing System Software
- Processing Framework (API)
- Science modules (processors)

3.4.2 Defining the operational system

In the discussion group, presentations were given by Jose Ramos, who presented a functional requirements diagrams describing a generic CCI system, and by Martin Boettcher (Landcover_cci) who discussed options for a centralised vs. de-centralised CCI system. His detailed presentation described how to incorporate continuous development in a grid network and included his suggestions for the best system option, which would be:

- Open and decentralized.
- Have a continuous algorithm improvement model: requirements, configuration management, EO data re/processing, validation.
- Enable continuous availability for repeated processing.
- Have common data access for climate modellers.
- Include new algorithms, configuration control of codes and environment used.
- Include a "science team" that could prioritize the development objectives.

The working group on the whole appreciated Martin's suggestions but it was suggested that the teams should be planning for a longer timescale and that in this model "cross-fertilization of ideas" is not promoted in itself.

3.4.3 Requirements for archiving and accessing data

The requirements for product dissemination need to be defined more clearly in order to develop a system that works efficiently. For example, if Cloud Computing were selected as part of the strategy, the system engineering teams would look for a-priori guidelines on how to set up the EO data archive within the Cloud. One question that would then arise would be how the archive should be implemented so that it is accessible through the Cloud, Grid or other configurations. Other points need to be defined for the harmonisation of data access before a system can be designed around it.

R2: The SEWG should start a discussion on how to archive data products for dissemination and whether this means further system requirements are needed.

3.4.4 Requirements must serve the scientific aims of the project

It is important not to introduce unnecessary barriers to development that could hinder the scientific work. Suggestions to encourage this were made:

- Phase 1 code to be evaluated with respect to a possible use in operational mode.
- System engineers are to review existing code and development plans to make them high quality, robust, good runtime performance, parallelizable, etc.
- System requirements have to be taken into account carefully.

3.4.5 Incorporating new elements

The teams must be aware of any future developments that could affect the operational system, whether these are changes to the input data, technological changes, scientific advances, etc. and the system should be designed with these in mind.

R3: The ECVs should consider how to incorporate future missions (especially the Sentinels) and whether doing so would result in a different set of requirements.

System flexibility and scalability is needed in defining operational environments to incorporate not only future missions also future requirements, to allow continuous algorithm improvement, to enable continuous availability for repeated processing and to have common data access for climate modellers.

Code portability must be ensured, as well as standards between different platforms and technologies. System developers need to be prepared to incorporate newer technologies as they arise.

The handling of network security for the operational system has not yet been discussed.

R4: The SEWG should explore security requirements from ESA, CCI projects' host organisations and other relevant organisations. The SEWG will derive a starting point to discuss across different systems.

3.5 Uncertainty Characterisation

The contributors for this chapter are:
Caroline Poulsen (Cloud_cci)
Kevin Tansey (Fire_cci)
Eva Haas (Soil moisture_cci)
David Tan (CMUG)
Nick Rayner (SST_cci)
Wouter Dorigo (Soil_moisture_cci)
Thierry Phulpin (CMUG)
Thomas von Clarmann (Ozone_cci)
Stephan Bojinski (WMO)
Simon Pinnock (ESA)

3.5.1 Aims

To obtain clarification on **what is required** by the users from the uncertainty characterisation and validation work in CCI.

To **discuss and share experiences** among the CCI teams on the approaches being undertaken in the characterisation and validation work, and to identify and try to resolve any common issues.

Uncertainty characterisation is identified as one of the strengths of the CCI but it's not clear that the projects are producing the necessary information that either CMUG or the users require.

3.5.2 Uncertainty characterisation

Teams should follow the uncertainty characterisation and validation guidelines produced at the first collocation meeting [ESA CCI Project Guidelines] but in addition:

- Product specification documents (PSDs) should be revisited in light of the uncertainty characterisation report.
- Products should include, at minimum, information on random and systematic uncertainties and if possible, correlated terms.
- PSDs and uncertainty characterisation documents should be reviewed by the internal climate research group, if it hasn't been already.
- Placeholders in the PSD are to be filled in later.
- Retrieval averaging kernel should be reported in the PSD.

R1: Products should include, at minimum, information on random and systematic uncertainties and if possible, correlated terms.

Quantitative and qualitative descriptions of the uncertainty estimations should be included in the uncertainty characterisation report. In some cases there are too many uncertainties to realistically address in the time given, here the ECVs should prioritise the uncertainties and analyse them accordingly. Meta-data for the products should have a link to the uncertainty characterisation report.

The full time-series should be validated as realistically as possible. The uncertainty estimates and their components should also be validated, for example if a bias correction is applied it should be reported in the product. An action might be to ask ECVs on their temporal validation protocol/plan.

R2: It is good practice to include uncertainties as part of the product and to validated the uncertainties in their own right. Projects should aim to validate every year of CCI data and make this part of any operational system.

The multi-dimensional maturity index (Bates & Barkstrom) was suggested as a useful tool for the projects as the development of the products can be tracked against the index. The process of error characterisation will make the products more mature. It could be part of the climate research groups' job to provide feedback on what makes an ECV's product immature and how it can be more mature. This way a small group can test the process first, and then CMUG could do the integrative work. Although there is a concern that within the final 6 months of the product the CRG will not have time to communicate their findings to CMUG.

However the validity of the Bates & Barkstrom maturity index was questioned for its applicability in assessing uncertainty. It was decided that before the Bates & Barkstrom maturity index can be used for this purpose it needs to be reviewed and updated.

3.5.3 Tools in uncertainty characterisation

Within these discussions the group identified where scope could be given to developing tools under the remit of uncertainty characterization. Tools could potentially be designed to:

- Re-bin the data and propagate the uncertainties correctly.
- Regionally and temporally average the time series and propagate the uncertainties.
- Monitor the stability of the time series.

SST_cci have already specified tools to address these issues.

3.5.4 Synergy across the projects

There are areas for synergy across the projects that should be explored, for example uncertainty in the Level 1b data and in calibration and uncertainty in ancillary information such as ECMWF fields. Information on the implemented version of these products should be documented.

Long-term consistency is a requirement, but is not being addressed by a number of ECVs. Many projects are optimised for a single year. These are challenging issues that need a lot of time but Phase 2 of the CCI will be too late to address these points. ECVs don't have to process the whole archive, only across the jumps, and should visit calibration and re-calibration (cal and re-cal could be a good topic for a future co-location meeting).

It would be useful if people working on the same instruments and doing the same processes could share information. Teams should investigate reports written by other projects and projects should make sure their reports are accessible on the web.

R3: Areas of synergy to be explored are Level 1b data – ancillary information and where projects are working on the same instruments/processes. Better communication across the projects is needed.

There was a concern that not all CCIs are aware of the uncertainties in the Level 1 input data.

Communicating uncertainty – ECVs that aren't doing this should learn from other CCIs on how they are communicating uncertainty to users, SST are a good example. Dedicated engagement of the climate research group is essential.

European Space Agency
Agence spatiale européenne

3.6 CCI Data Products, Standards and Access

The contributors for this chapter are:
Ralf Quast (SST_cci)
Martin Stengel (Cloud_cci)
Max Schwinger (Aerosol_cci)
Max Reuter (GHG_cci)
Arnd Berns-Silva (Fire_cci)
Paul van der Linden (CMUG)
Martin Juckes (STFC/RAL)
Victoria Bennett (ESA)
Craig Donlon (ESA)

3.6.1 Aims and first steps

The discussion group and drafting group have the following aims:

- Produce a plan for good data management plan, which will save a lot of effort and cost later.
- Ensure a concerted effort on data management to obtain agreement on common issues, the priorities being to simplify user access across all projects and to simplify data management

The first steps taken towards these aims have been to review the Data Standards Working Group (DSWG) best practices document ('Guidelines for data producers'), to which no objections have been raised. The working group has also answered 7 questions posed to the meeting.

The teams have agreed to provide sample data by the end of January 2012 in the form of single files meeting data syntax specifications (but not necessarily scientific accuracy objectives).

3.6.2 How should ECV products be archived and disseminated to maximize uptake for climate science?

NetCDF & CF conventions have been accepted by consensus. Shape and GeoTIFF files may be used in addition (although lack of standards equivalent to CF means meta-data will be less robust). To make data management workable, NetCDF should be the baseline, with the above meta-data standards. The SST_cci PSD (also Ocean Colour and Sea Level_cci) provides a useful starting point, based on 12 years of GHRSST experience.

R1: There should be CF standard names for all variables (teams shall check and register names now, if needed) and also consistent, cross-CCI documented vocabularies and acronyms for platform, sensor, instrument, institute etc.; these should be listed on the website.

Advice should be taken from CMIP5/CF community on specifying spatial grids.

3.6.3 How do we ensure the consistency of ECV products?

Tools should be provided for checking compliancy with data requirements and 'discovery' meta-data and for generating compliant files. A standard file-naming recipe will be adopted and the teams should agree on the usage of 'licence' and 'publisher' attributes.

R2: For baseline data requirements, files should all contain the following (more information in 'Guidelines for data producers): a unique tracking ID; URLs for documentation; title, institution, history, source, reference, conventions, product version.

3.6.4 Do we need consistent data access policies across all projects?

Various protocols and approaches for accessing data were discussed, including FTP (anonymous), and the THREDDS data server package of services.

Suggestions for unified directory structures (details to be confirmed) could look like:

- YYYY/MM/dd/inst/sensor/level/version/variable
- inst/sensor/level/version/variable/YY/MM/dd (using hard and symbolic links)

Compression of data will be useful for some but not appropriate for all data. The data should be managed locally and copied to central ESA archive. Teams wanting to provide data access to their data should do so through anonymous FTP servers. One or more central archive(s) will collect data and serve it through a THREDDS data server. There is a need to maintain synchronization, to use a Digital Object Identifier and provide a discovery portal and tools. The central archive(s) should run an Earth System Grid Federation Data Node.

3.6.5 What should be the next steps for the data standards working group?

- Consolidate data standards vision and structure.
- Collect (by Jan 2012) and review demonstration products.
- Hold (at least) one meeting and regular tele-conferences (3 per year).

3.6.6 Outstanding issues

R3: The projects should produce an accessible log file that will record the key steps in the processing, to be generated by the processing software as well as which input data products have been used (e.g. accessible via URL in data file) and/or a list sources (syntax to be agreed).

The terms of use for the data products must be: harmonised, open for commercial use, subject to normal citation requirements and allow redistribution.

In terms of data visibility, the data must be search engine-friendly and linked to from relevant portals. The projects need to consider whether additional meta-data (i.e. structured documentation not contained in files) is needed/will be useful.

3.7 Data and Model Confrontation

The contributors for this chapter are (alphabetic order):
Alexander Loew (CMUG),
Stephen Plummer (ESA)
Mark Ringer (CMUG),
Michael Schulz (Aerosol_cci),
Peter van Velthoven (Ozone_cci)

3.7.1 Objectives

The discussion focussed mainly on the use of observational data sets for climate modelling but also considered some issues related to the application of the data more widely in climate science.

3.7.2 Status

High quality climate observations are essential to (a) ensure reliable assessments of the quality of climate model simulations and (b) contribute to their continuing development and improvement. Observational data is used in several ways:

1. Benchmarking and evaluation of model simulations
    - Objective ranking of models, including the development of reliable performance metrics
    - Assessing if models are fit for purpose
    - Evaluating the impact of changes made during the model development cycle
2. Process studies
    - Better process understanding and description of physical processes
    - Improved parameterization/representation of key physical processes
3. Initial and boundary conditions
    - Seasonal-to-decadal predictions
    - Re-analyses
    - Certain types of simulation (e.g. atmospheric models forced by observed SSTs and sea-ice.

Using better observations will drive further model improvement in these various application domains. The diversity of these applications requires different datasets. **A single dataset is highly unlikely to fit all requirements.** Model evaluation and thus the required observations cover a wide range of spatial and temporal scales.

Overarching requirements from a model user perspective are long-term harmonized L2 datasets based on robust L1 best calibrated data documented competitive high quality of data consistency between data sets where appropriate (see below).

Error characterization: observational errors are needed to determine if improved models are actually getting closer to the observations and to judge if an ensemble of models covers the full range of uncertainty. The panel recognizes that it is of particular importance for CCI to achieve a high level of:

- Visibility of the data in the community. It is recognized that there are still deficits in the recognition of current ESA data in the user community (e.g. iLAMB).
- Easy accessibility of the data for the user community. It is crucial that the data will actually be used in the end. Easy access and concise but clear documentation is considered to be essential
- Future proofing. It is very important for the CCI to recognize that models are developing quickly and in many ways (e.g. increasing resolution, inclusion of new physical processes, assimilation) and to try and keep pace with and anticipate these developments. This is essential to ensure that the data sets do not become over-simplistic in a relatively short time.

- The JPL-led initiative for providing observational data sets for CMIP5 analysis is a very good example.

R1: The CCI projects should envisage data delivery to an ESG-CMIP5 data portal e.g. at BADC or MPIM in CF compliant format using CMOR routines and recommended CMIP5 data publishing practices to make a contribution to the analysis of CMIP5 simulations. The panel recognizes that there is a good opportunity for CCI to contribute to the scientific analysis of CMIP5 in the next 5-10 years even before AR6.

3.7.3 Product Uncertainties

One of the major assets of ESA CCI is that all the projects will provide quantitative uncertainty estimates to go with their products. While the use of this uncertainty information seems to be straightforward for some applications, e.g. re-analysis and weighting of observations by their uncertainties, it still remains unclear how uncertainty information might be actually be used in practise by the different users in the consortia and the wider user community. To ensure that the uncertainty information is appropriate and that it will actually be used and that CCI efforts are spent efficiently, the panel recommends that the CCI, as a whole, conducts a survey of the users.

R2: A survey of the use of uncertainty information by the users (What do they mean by uncertainties? Why do they require them? How will they use this information?). This is to be done by the CMUG for the climate modelling community and by the CCI teams (through their respective climate research groups) for other applications.

3.7.4 ECV data analysis

The panel emphasizes the need for an early and comprehensive user feedback in Phase 2 of the CCI.

R3: A first version of the long-term ESA CCI ECV products should be available to the users at end of the first year of Phase 2 of the programme.

R4: Appropriate user feedback is ensured by further strengthening the involvement of users in Phase 2. This should initially help to generate momentum in the various user communities and subsequently ensure their continued engagement with the CCI.

R5: A strong commitment is needed to regular reprocessing cycle at appropriate times to feed into international activities (e.g. CMIP-XX).

R6: Emphasise scientific exploitation and the further involvement of scientific users in the programme throughout Phase 2 (and beyond). The latter should include science users that are not involved in climate modelling, e.g. those involved in trend studies.

3.7.5 Enhancing visibility and data usage

To enhance the visibility of CCI datasets, the panel recommends:

R7: Quick-look documentation: Provide for each data set a Summary for Scientific Users (SSU) that concisely documents the characteristics of each dataset and its application limits. The panel highly recommends that CCI products follow the template applied successfully for CMIP5 (see PCMDI website) by JPL and others.

R8: A clear documentation of the added value of the data products (included in the SSU) is required in order for users to make reliable assessments of their utility.

3.7.6 Consistency

The panel recognizes the importance of consistency between different ECV data products in terms of temporal and spatial coverage, assumptions, use of ancillary data and internal logic. It is recognized that it is not necessary that all of the ECV products envisage consistency with all of the others. Depending on the importance of the links between ECV's for specific user applications, more or less emphasis has to be given to consistency. In some cases it may also be useful to follow conventions applied in external datasets to facilitate comparison and/or synergistic use.

R9: CMUG should consult with the CCI-teams to provide an analysis of the importance of consistency between different ECVs from a climate modelling perspective (i.e. in what cases is it of greater or lesser importance?)

Inconsistencies between different data sets is considered to be of importance both scientifically but also from a 'lessons learnt' perspective.

R10: Document and communicate inconsistencies between different data sets in data products.

3.7.7 Golden year

The panel discussed whether there is actually a need for a single golden year. A pragmatic approach might be to group different ECV's in accordance to their logical relationship (e.g. landcover, fire, aerosols, GHG or SST and sea ice or clouds and aerosols) and aim for a golden year at least for these groups.

Subsequently to these discussions, CMUG was tasked with undertaking a consultation of the projects to decide the best dates for a golden year. The year chosen was 2008, which suits most projects.

3.8 CCI System Requirements

The panelists were:
Gilles Larnicol (Sea-Level_cci)
Mark Dowell (CEOS – WGClimate)
Thomas Holzer-Popp (Aerosol_cci)
Chris Merchant (SST_cci)
Carsten Brockmann (Ocean_Colour_cci and Land_Cover_cci)
René Forsberg (Ice_Sheets_cci)

The chairs were:
Pascal Lecomte (ESA)
Mark Doherty (ESA)

The contributors for this chapter are:
Paul Spinks (SST_cci)
Thomas Nagler (Glaciers_cci)
Máximo Fernándes Cortizo (Fire_cci)
Christina Aas (Ice_sheets_cci)
Thomas Holzer-Popp (Aerosol_cci)
Harmut Boesch (GHG_cci)

As a large part of the collocation meeting was dedicated to planning for Phase 2, there was both a panel debate and a drafting group on CCI system requirements. The debate began with some presentations from the system engineering groups before broad discussions on whether there are universal requirements across the projects for quality, data standards and data access. Consistency, as required by GCOS, depends on inputs so system requirements on input data could also be identified. [Sections 3.8.1 – 3.8.6]

Questions the audience were asked were: how can cross-fertilisation between ECVs be encouraged, and what would be the trade-off if sharing or not sharing infrastructure between ECVs?

Where there is uncertainty on how often to update and regenerate the records, CMUG, i.e. user feedback should be driving the system. Many projects are working on legacy systems.

3.8.1 What are the performances requirements for the products (in terms of quality) for Phase 2, i.e. the next 3 years of the CCI? Consistency, accuracy, stability.

The main goal within the first year is to produce data with documented quality and uncertainty, as users are waiting for this. After 3 years the projects should be aiming for a full reprocessing of the data – assuring best quality, repeated algorithm improvement, reprocessing.

After 3 years it would be great if the CCI could embed their data streams within active climate services and make the data streams indispensable.

Need a two-way interaction with users about uncertainties, need to communicate how to use them.

Performance with respect to consistency is needed. From a system perspective it's key to provide a framework that ensures consistency across the CCI. In the long-term, improvements in consistency are needed, in particular by the end of Phase 2. For further continuation of the ECVs (post CCI) centralised funding may not be required; long-term sea-ice projects have been run with local/national funding for international efforts. E.g. TBC – geology, gravity service, run by the French for 50 years. Should be preparing to provide a long-term service run on trans-national cooperation.

Stability/accuracy in each ECV needs to be tackled by science leader – the system part is different.

Sea level has had a hard time to reach user requirements anyway so for defining Phase 2 ambitions Sea Level will need a specific way of improving the error characterization, using the round robin to find where the errors are. This with user feedback can show the benefit of the data.

The CCI should use a maturity index to track the progress of a product and see how it's developing. It will be useful to develop indicators of how much progress ECVs are making through phases 1-3. Maturity will vary between ECVs at present.

Main focus for new projects is different, i.e. improving the accuracy. On a longer time scale it should be on making a continuous time series available.

Traceability is important for the future. Must keep all the elements of meta-data with the datasets – which algorithms used, which sensors it's from, etc. Also establish international links e.g. to WGClimate. It is important at level 1 that there is consistency across missions.

3.8.2 Performance objectives for Phase 2 – is it important to deliver consistent 13 ECVs for AR6?

In short, yes.

3.8.3 Should the projects be syncing their reprocessed datasets to be aligned with the IPCC deadlines? Or should they be independent of it?

The AR cycle is 5-6 years which is too long for a reprocessing cycle; having said this, the projects don't want to be doing it too often – SST users say a reassessment every 2 years is a good length. The projects should be able to deliver final datasets every 2 years but revisit the algorithms on a different timescale. The internal cycle is more frequent than 2 years. The projects need to be able to revisit and update their datasets within a timescale of weeks otherwise the whole process becomes unfeasible.
To be used in AR6 the projects must produce a dataset of 10 years minimum, with a full time series even better. The CCI should not rely on just putting the data 'out there' and hoping they get used; we need to showcase good examples with groups of users to demonstrate the usefulness of the datasets.

Some projects do intend to publish work in time for AR5.

Sea Level has a consistent sea level data set of about 20 years' length but it should also produce regular products on monthly basis – currently they reprocess every 2 years (in full), but continue to follow the mean SL trends so the project ends up with a long time series and a current status.

3.8.4 What is the cost of the 'system' to operationally produce 30 years (and more) of 13 ECVs, in terms of organisation, implementation and operation?

It's too early to make any firm decisions, as the projects didn't know what was in the system until recently, the ECV statuses have been too different and the system needs to be defined better. Martin Boettcher showed different scenarios in his presentation, the projects could do the same for different ambitions CCI has for the system.

There are two main options – decentralised or not? It depends which resources you need, which would then affect the cost. If scientists do the operational work they are not available for research.

A first estimate on cost, starting from the current project but increasing the system engineering part and downscaling the algorithms to a smaller section, using the current budget to estimate, would say €10m per annum. For optimum cost efficiency the CCI could look at all the processes and see what can be shared between projects, e.g. level 1 data, documentation and promotion.

Taking MyOcean as an example can give some idea of the cost – this has €33million for 3 years from the EC but a total cost €50m, which takes into account service and transverse activities. The cost of around €10M/year for 12 geophysical parameters (a similar number to the current CCI programme) and production centres is a feasible number and similar to the first estimate above.

If CCI follow the first estimate costings, then the up-scaling approach also misses the dissemination costs. In the development phase we would expect these to be a considerable part. Need to be careful about using current costs for future estimates. Although the CCI will be taking advantage of current systems developed with a lot of previous investments, building the system will require more money than the sum of the individual parts– there will be additional costs.

Level 1b inputs can be shared, and some common Level 1 issues can be tackled, such as problems with input data (this would be an additional cost as we don't look at it now). Multisensor data for validation can do more than just validate (see presentation yesterday), it is also a way to cross between ECVs to perform a consistency analysis. Could take in other products alongside SST to address inconsistencies (but this is an additional cost too). Level 1 issues are relevant but these costs shouldn't come under the CCI, they are the responsibility of the space agencies.

3.8.5 Are 13 SRDs and associated documents sufficient to start Phase 2 or do we need a higher-level requirement document to describe the common parts… or not?

Yes, it is needed as the 13 SRDs are driven by individual experiences; it could be produced now or in Phase 2 using Phase 1 documents as input, perhaps summarising what has been learnt in the first year. Similarly if we wanted a distributed but integrated system then we need transverse system commentary; it would be useful to have a couple of scenarios for decentralised and centralised ideas. It could also be used to address shortcomings, e.g. the need in Ice Sheets to supplement the ECV with other 3rd party data.

The panel agree it is a good idea but are wary for signing up to write more documents!

In M Doherty's slides, international activities were shown with a top down vision but this system perspective from CCI will be bottom up. There are differences and commonalities in both system designs – if you can extract these, they will be useful for activities doing similar things on an international level, i.e. other space agencies.

3.8.6 How do you see the CCI system after 2016?

It should not be visibly or scientifically disaggregated, it needs to be visible and well linked to climate. It could have a choice of long-term climate records but also GMES NRT focus. Would prefer a data record focus. In GMES there are downstream services – which NRT seems to be but it seems to have a different user base, which implies more cost. Although other ECVs wouldn't make distinction between the two time scales, e.g. for SST they are consistent. It is important that each ECV should be able to provide a regular time series and also a short (monthly) current state (as e.g. for MyOcean).

Unique access could be good for users (on distributed system), or be specific for ECV.

The overall vision for 2016 is for the ECVs to still exist and to be used! Therefore it must be able to evolve to meet changing requirements. Originally the different parts of ECV production have been quite separate but this is now changing, instead of 3 sets of people, the operations, scientists and users will merge. Maybe after 2016 the aim should be to deliver more than files? The users should be more integrated in the whole process.

Thinking more into the future, a new entity could exist that oversees and provides this system, e.g. a climate observing system, similar to weather systems. If the climate services need is strong enough and generates strong external requirements then this is a possibility.

It is important that the products are upgraded and new elements are added, including requirements coming through users. Having a bottom up process drives this.

In the drafting team the group brought together some of the elements from the panel debate with new discussions to come up with requirements for the CCI:

3.8.7 Stakeholders

User requirements exist at ECV level but not at the CCI programme level; the drafting group looked at what statements can be made about user requirements at a higher level:

- Who are the users?
- What combinations of ECVs do they want?
- How many users need each combination?
- Which (types of) users are core users?

The projects will need to bear in mind that algorithm developers and end users will have different needs but the climate research group and CMUG represent a lot of the science user groups.

R1: Algorithm developers, scientists and end users should be involved in drafting CCI system requirements.

3.8.8 High level system requirements

The results of this discussion should form the basis for CCI Phase 2. Currently there is no adequate process for identifying high-level system requirements. The set of ECV SRDs may contain gaps or inconsistencies; collocation meetings are too short to address such issues in sufficient details.

R2: ESA should initiate a process for capturing high-level CCI system requirements.

Inputs should include: URDs and SRDs, drafting notes from Colocation 2 splinter groups, GCOS and other external requirements.

Science leaders and system engineers could draft this, to be reviewed by the rest of the project team.

3.8.9 Continuous improvement processing

Users need assurance that data are being curated so that it is guaranteed they are stable, up to date (extending observations backwards and well as forwards), relevant/contemporary and quality controlled.
Most end users need controlled updates at regular intervals, as well as rapid responses in the event of problems.

R3: For production systems, a continuous improvement process should be implemented.

Algorithm developers need agile, responsive systems to allow a combination of stability and easy, rapid development. At the moment there isn't an agreed period for the update cycle so it could be designed to synchronise with other programmes (e.g. IPCC). See Figure 1.

R4: Productions systems should ensure availability of older versions of the data (although avoiding confusion over too many versions).

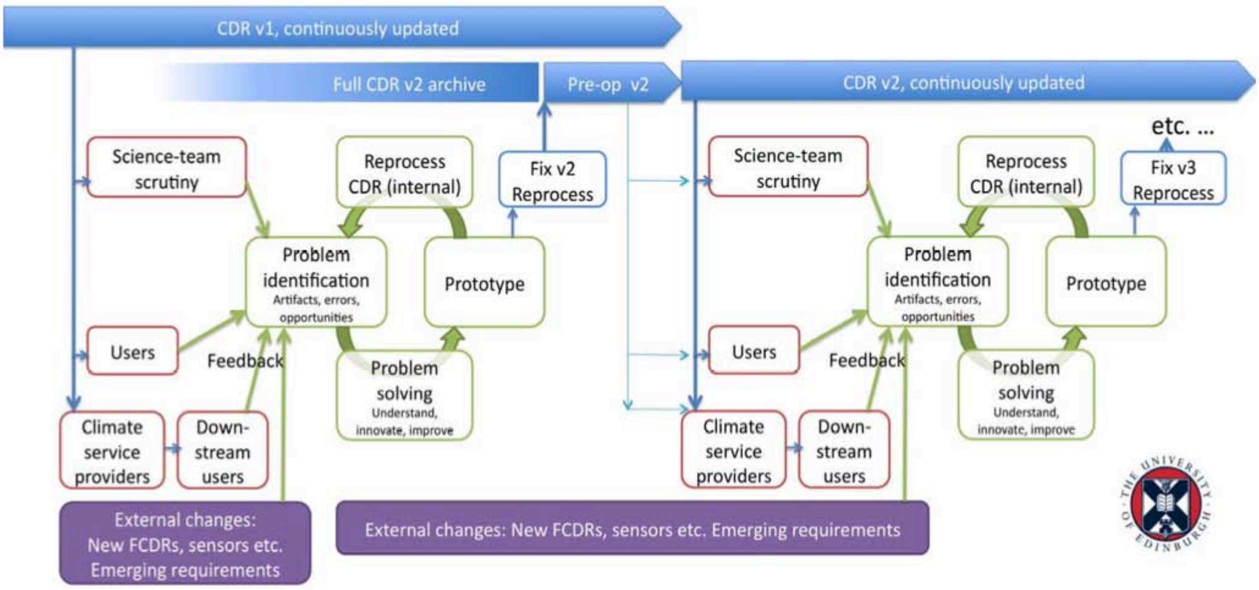

Figure 1. Climate data record continuous improvement process schedule from SST.

3.8.10 Interdependencies between the ECVs

Little or no support is given to the reuse of code between ECV systems. More support could be given to ensure the use of common codes, based on consistent use of Level 1B data, e.g. using Aerosol ECV for atmospheric correction.

R5: Encourage the reuse of code between ECVs. This could be achieved by greater use of open source algorithms.

3.8.11 Functional architecture

Most users shouldn't care about system architecture as long as it meets their needs; algorithm developers care more about the architecture.

Should the interface with users be consistent across all ECVs or should community norms be adopted? Some common elements could be identified, e.g. archiving, data dissemination. See Figure 2.

R6: ESA should adopt a CCI functional architecture to facilitate further discussions.

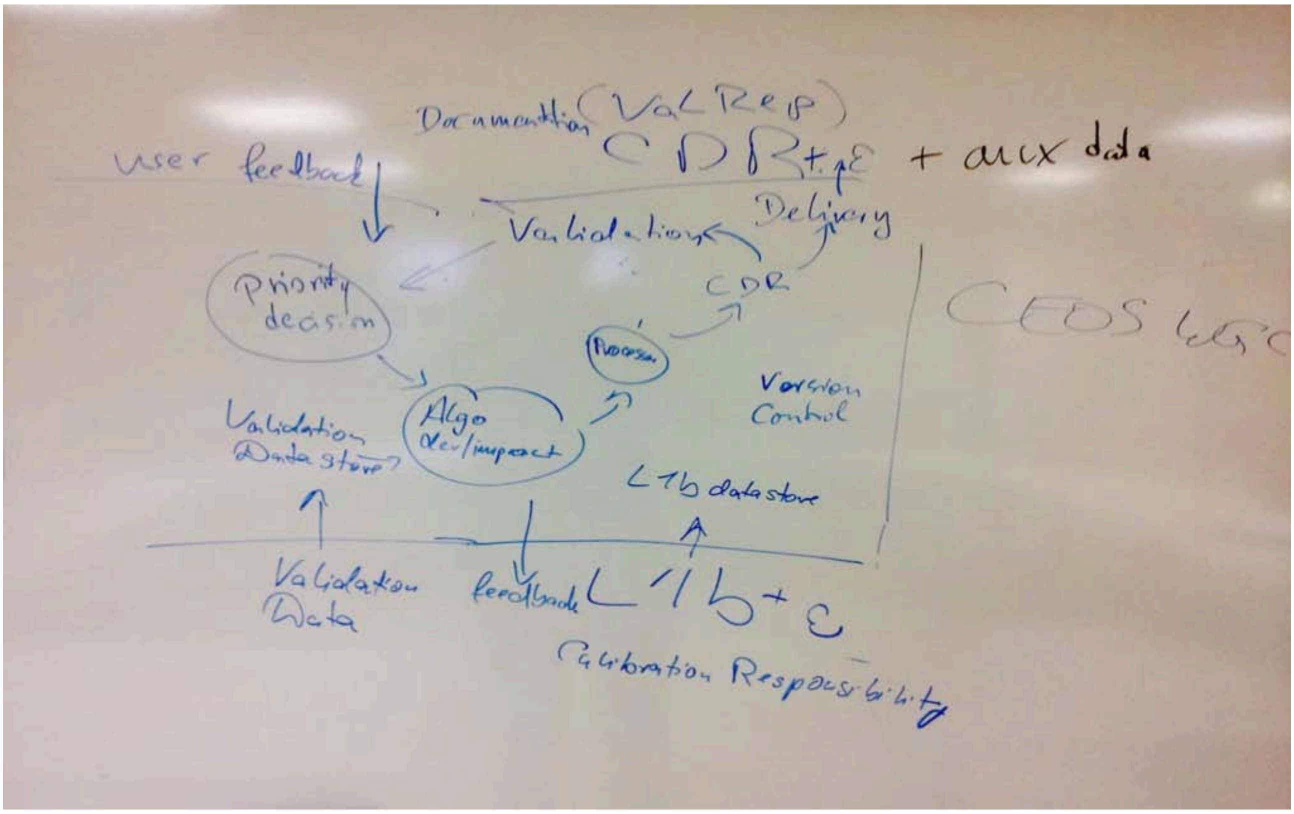

Figure 2. Functional architecture suggestion from the system requirements drafting team.

APPENDIX A RECOMMENDATIONS SUMMARY

Science Agenda	
SA-1	A community paper is to be written by the science leaders; a joint CCI session should be organised at an international conference; the CCI data sets are to be made available alongside international climate data archives (e.g. CMIP5/ESGF); a special issue of a journal publishing CCI Round Robin results is to be investigated.
SA-2	ESA will coordinate the production of a CCI brochure and an iPad app, with input from the projects.
SA-3	Expand the CCI Science Agenda developed at the last collocation to develop key scientific questions that exploit the data products and involve multiple ECVs.
SA-4	The next CMUG Integration meeting should have dedicated time for the science leaders to get together to focus on the Science Agenda. They should have more input in setting the agenda for the meeting.

Algorithms and Performances	
AP-1	The round-robin and selection inter-comparison should continue throughout the CCI programme to include periodic re-evaluations of algorithms, adoptions of new algorithms and re-processing of data archive, as and when is necessary.
AP-2	Maintain traceability in the round-robin planning and execution.

Data Delivery, Quality and Feedback	
DQF-1	As recommended by CEOS, space agencies should maintain a dedicated website for routine sensor monitoring, data quality reports (plus summary) and uncertainty characterisation. The site should contain complete knowledge of the sensor, platform and data processing so that the information is traceable, as recommended in QA4EO.
DQF-2	All ECV data products shall contain quality indicators as well as information on uncertainties. In addition, all products should have a clearly defined versioning system and should have a unique DOI for traceability.
DQF-3	All ECV data products should be easily available through a common portal that supports (a) multiple download mechanisms (e.g. FTP/HTTP and appropriate media for large volumes), (b) subscription services and (c) product sub-setting.
DQF-4	ESA should consider establishing a User Processing Facility (UPF) for rapid processing and reprocessing of ECV data products. The system would allow users to login remotely and run their own software on the UPF.

DQF-5	A routine mechanism should be established for feedback between CCI project teams and QWGs responsible for ESA and 3rd party mission Level 1b data products.
DQF-6	All CCI project teams need to regularly exchange information on data quality.

System Engineering	
SE-1	The SEWG should review requirements and analyse requirements across the CCI to derive commonalities from the preliminary system analyses.
SE-2	The SEWG should start a discussion on how to archive data products for dissemination and whether this means further system requirements are needed.
SE-3	The ECVs should consider how to incorporate future missions (especially the Sentinels) and whether doing so would result in a different set of requirements
SE-4	The SEWG should explore security requirements from ESA, CCI projects' host organisations and other relevant organisations. The SEWG will derive a starting point to discuss across different systems.

Uncertainty Characterisation	
UC-1	Products should include, at minimum, information on random and systematic uncertainties and if possible, correlated terms.
UC-2	It is good practice to include uncertainties as part of the product and to validated the uncertainties in their own right. Projects should aim to validate every year of CCI data and make this part of any operational system.
UC-3	Areas of synergy to be explored are Level 1b data – ancillary information and where projects are working on the same instruments/processes. Better communication across the projects is needed.

Data Products, Standards and Access	
PSA-1	There should be CF standard names for all variables (teams shall check and register names now, if needed) and also consistent, cross-CCI documented vocabularies and acronyms for platform, sensor, instrument, institute etc.; these should be listed on the website.
PSA-2	For baseline data requirements, files should all contain the following (more information in 'Guidelines for data producers): a unique tracking ID; URLs for documentation; title, institution, history, source, reference, conventions, product version.
PSA-3	The projects should produce an accessible log file that will record the key steps in the processing, to be generated by the processing software as well as which input data products have been used (e.g. accessible via URL in data file) and/or a list sources (syntax to be agreed).

Data and Model Confrontation	
DM-1	The CCI projects should envisage data delivery to an ESG-CMIP5 data portal e.g. at BADC or MPIM in CF compliant format using CMOR routines and recommended CMIP5 data publishing practices to make a contribution to the analysis of CMIP5 simulations. The panel recognizes that there is a good opportunity for CCI to contribute to the scientific analysis of CMIP5 in the next 5-10 years even before AR6.
DM-2	A survey of the use of uncertainty information by the users (What do they mean by uncertainties? Why do they require them? How will they use this information?). This is to be done by the CMUG for the climate modelling community and by the CCI teams (through their respective climate research groups) for other applications.
DM-3	A first version of the long-term ESA CCI ECV products should be available to the users at end of the first year of Phase 2 of the programme.
DM-4	Appropriate user feedback is ensured by further strengthening the involvement of users in Phase 2. This should initially help to generate momentum in the various user communities and subsequently ensure their continued engagement with the CCI.
DM-5	A strong commitment is needed to regular reprocessing cycle at appropriate times to feed into international activities (e.g. CMIP-XX).
DM-6	Emphasise scientific exploitation and the further involvement of scientific users in the programme throughout Phase 2 (and beyond). The latter should include science users that are not involved in climate modelling, e.g. those involved in trend studies.
DM-7	Quick-look documentation: Provide for each data set a Summary for Scientific Users (SSU) that concisely documents the characteristics of each dataset and its application limits. The panel highly recommends that CCI products follow the template applied successfully for CMIP5 (see PCMDI website) by JPL and others.
DM-8	A clear documentation of the added value of the data products (included in the SSU) is required in order for users to make reliable assessments of their utility.
DM-9	CMUG should consult with the CCI-teams to provide an analysis of the importance of consistency between different ECVs from a climate modelling perspective (i.e. in what cases is it of greater or lesser importance?)
DM-10	Document and communicate inconsistencies between different data sets in data products

System Requirements	
SR-1	Algorithm developers, scientists and end users should be involved in drafting CCI system requirements.
SR-2	ESA should initiate a process for capturing high-level CCI system requirements.

SR-3	For production systems, a continuous improvement process should be implemented.
SR-4	Productions systems should ensure availability of older versions of the data (although avoiding confusion over too many versions).
SR-5	Encourage the reuse of code between ECVs. This could be achieved by greater use of open source algorithms.
SR-6	ESA should adopt a CCI functional architecture to facilitate further discussions.

APPENDIX B LIST OF PARTICIPANTS

		Panels			Working Groups						Drafting Teams							
		CCI Project Status	Key Sciences Issues	CCI System Requirements	Algorithm and Performances	Data Delivery, Quality & Feedback	System Engineering	Uncertainty Charact. & Validation	Data Standards & ECV Prod. Access	Data & Model confrontation	Key Sciences Issues	Algorithm and Performances	Data Delivery, Quality & Feedback	System Engineering	Uncertainty Charact. & Validation	Data Standards & ECV Prod. Access	Data & Model confrontation	CCI System Requirements
Cloud	Rainer Hollmann																	
	Martin Stengel																	
	Caroline Poulsen																	
Ozone	Michel Van Roozendael																	
	Ronald Van der A																	
	Meinhard Wolfmueller																	
	Thomas von Clarmann																	
	Peter van Velthoven																	
Aerosol	Thomas Holzer-Popp																	
	Gerrit de Leeuw																	
	Max Schwinger																	
	Michael Schulz																	
GHG	Michael Buchwitz																	
	Hartmut Boesch																	
	Max Reuter																	
	Günther Lichtenberg																	
Sea-Ice	Stein Sandven																	
	Geoff Busswell																	
	Stephan Kern																	
Sea-Level	Anny Cazenave																	
	Gilles Larnicol																	
	Detlef Stammer																	
	Geoff Busswell																	
SST	Chris Merchant																	
	Paul Spinks																	
	Gary Corlett																	
	Nick Rayner																	
	Ralf Quast																	

European Space Agency
Agence spatiale européenne

		Panels			Working Groups							Drafting Teams							
		CCI Project Status	Key Sciences Issues	CCI System Requirements	Algorithm and Performances	Data Delivery, Quality & Feedback	System Engineering	Uncertainty Charact. & Validation	Data Standards & ECV Prod. Access	Data & Model confrontation	Key Sciences Issues	Algorithm and Performances	Data Delivery, Quality & Feedback	System Engineering	Uncertainty Charact. & Validation	Data Standards & ECV Prod. Access	Data & Model confrontation	CCI System Requirements
Ocean-Colour	Shubha Sathyendranath																	
	John Swinton																	
	Bob Brewin									?								
	Mike Grant																	
	Carsten Brockmann																	
Glaciers	Frank Paul																	
	Tobias Bolch																	
	Thomas Nagler																	
	Tazio Strozzi																	
Land cover	Pierre Defourny																	
	Sophie Bontemps																	
	Carsten Brockmann																	
	Martin Boettcher																	
Fire	Emilio Chuvieco																	
	Kevin Tansey																	
	Martin Bachmann																	
	Oscar Navarro																	
	Máximo Fernández Cortizo																	
	Arnd Berns-Silva																	
Soil Moisture	Wolfgang Wagner																	
	Eva Haas																	
	Martin Ertl																	
	Wouter Dorigo																	
Ice Sheet	René Forsberg																	
	Christina Aas																	
	Thomas Nagler																	
CMUG	Roger Saunders																	
	Mark Ringer																	
	Alexander Loew																	
	David Tan																	
	Paul van der Linden																	
	Thierry Phulpin																	

European Space Agency
Agence spatiale européenne